

Metodiky ontologického inženýrství

Helena Kučerová

Vyšší odborná škola informačních služeb, Pacovská 350/4, 140 00 Praha 4

Postprint původní studie:

KUČEROVÁ, Helena. Metodiky ontologického inženýrství. *Ikaros* [online]. 2011, roč. 15, č. 5 [cit. 31. 05. 2011]. Dostupné z: <http://www.ikaros.cz/node/6801>. URN-NBN:cz-ik6801. ISSN 1212-5075

Abstrakt

Článek je rozdělen na dvě základní části. První část shrnuje současný stav vývoje metodik ontologického inženýrství a nabízí jejich typologii. Ve druhé části je prezentováno sedm konkrétních metodik, jež byly vybrány s cílem představit co nejrozmanitější přístupy k návrhu ontologií. Popis metodik má jednotnou strukturu, tak aby umožnil jejich srovnání. Je zaměřen především na procesy realizované v průběhu návrhu a na jejich vzájemné vztahy.

Klíčová slova:

metodiky, ontologické inženýrství, ontologie, workflow, životní cyklus projektu

1. Úvod

Termín ontologie byl donedávna spojován především s filozofickou disciplínou zkoumající obecné principy bytí. V poslední době s nástupem sémantického webu se tento termín v poněkud posunutém významu čím dál častěji objevuje i ve slovníku knihovníků a informačních pracovníků. Za poměrně krátkou dobu zhruba od 90. let 20. století už informační teorie i praxe stačila vyvinout značné množství často velmi rozdílných koncepcí „informatických“ ontologií. Tuto šíři pojetí dokládá například ontologický metamodel konsorcia OMG, který nabízí následující vymezení obsahu a zejména rozsahu pojmu ontologie: „Ontologie definuje obecné termíny a pojmy (významy), používané k popisu a reprezentaci určité oblasti vědění. *Ontologie může co do expresivity sahát od taxonomie (znalosti strukturované hierarchicky nebo genericky stylem předeek – potomek) k tezaurům (slova a synonyma), konceptuálním modelům (se složitějším zachycením struktury znalostí) až k logické teorii (s velmi bohatými, složitými, konzistentními a smysluplnými znalostmi).*”¹

Knihovníkům a informačním pracovníkům zní takový výčet velmi povědomě – je pro ně důkazem, že ontologie se nevynořují odkudsi z neznáma, ale naopak že navazují na dosavadní snahy odborníků i laiků organizovat informace a znalosti, ať už byly příslušné činnosti nazývány kategorizace, třídění, klasifikace, indexování, věcná katalogizace nebo věcné pořádání, a jejich výsledky označovány jako indexy, katalogy, klasifikace, číselníky, autority, hesláře či selekční jazyky. Co do formy bývají ontologie konstruovány jako pojmové hierarchie nebo sítě – opět známý koncept z historie organizace znalostí.

Ontologii chápeme pro účely tohoto textu jako *sdílitelnou a opakovaně použitelnou pojmovou reprezentaci vymezené domény* (úseku reality). S trochou zjednodušení lze říci, že zatímco cílem slovníků je nastolit pořádek mezi slovy a klasifikace se o totéž snaží u pojmů, vlastním zájmem ontologie je realita², přesněji řečeno znalosti o ní. Na rozdíl od filozofické ontologie však není cílem informatické ontologie vytvořit reprezentaci reality v její úplnosti, zaměřuje se vždy na určitou dílčí

oblast, vymezenou tzv. univerzem diskursu (oborem úvah)³. Je dokonce možné vytvořit více ontologií pro jeden výsek reality, pokud na něj pohlížíme z různých úhlů pohledu a zaměříme se na rozdílné typy řešených problémů a úloh (např. jinou ontologii rostlin si zřejmě vytvoří zahrádkáři a jinou biogenetici). Za hlavní oblasti využití ontologií se považují organizace a vyhledávání znalostí, komunikace znalostí, opakované využití existujících znalostí a automatické odvozování (inference) nových znalostí. Tohoto širšího záběru oproti tradičním systémům organizace znalostí dosahují ontologie díky bohatší a jemnější propracované sémantice, zejména v oblasti vztahů, axiomů a pravidel.

Dosavadní vývoj naznačuje, že v distribuovaném síťovém prostředí webu se nepředpokládá vytvoření jediné všeobsahující univerzální ontologie⁴, naopak jsou podporovány doménově a úlohově individualizované ontologie navrhované na míru potřebám jednotlivých specifických komunit. S jistou nadsázkou by bylo možné parafrázovat druhý Ranganathanův zákon knihovní vědy ve znění „každému uživateli jeho ontologii“. Je tedy pravděpodobné, že mnozí informační profesionálové budou v dohledné době postaveni před úkol navrhnout vlastní ontologii.

Následující přehled metodik ontologického inženýrství si vzhledem k explozi publikací věnovaných ontologiím v posledním desetiletí samozřejmě nemůže dělat nároky na úplnost. V současné době už existuje nejen množství metodik, ale byly již vypracovány i četné studie zaměřené na jejich srovnávání. Jim nechce tento text konkurovat. Snažíme se spíše svým výběrem představit současnou rozmanitost přístupů a případné zájemce o další metodiky odkazujeme na komparativní zdroje uvedené v závěrečném seznamu literatury. Naším cílem není ani vyhodnotit „nejlepší“ metodiku – pro tento účel je možné využít např. vícekriteriální systém hodnocení a výběru metodik METES (Methodology Evaluation System), navržený Alenou Buchalceovou původně pro oblast softwarového inženýrství⁵. Tak jako existují různé typy ontologií, existuje i velké množství různě zaměřených metodik, výběr té nejvhodnější pak ovlivňují především specifika věcné oblasti a typu navrhované ontologie. V charakteristikách jednotlivých metodik se pokusíme ukázat jejich užitečnost pro specifické účely a pomoci tak zájemcům o tvorbu ontologií vybrat si takový postup, který bude vhodný pro jimi realizovaný projekt.

2. Ontologické inženýrství a metodiky

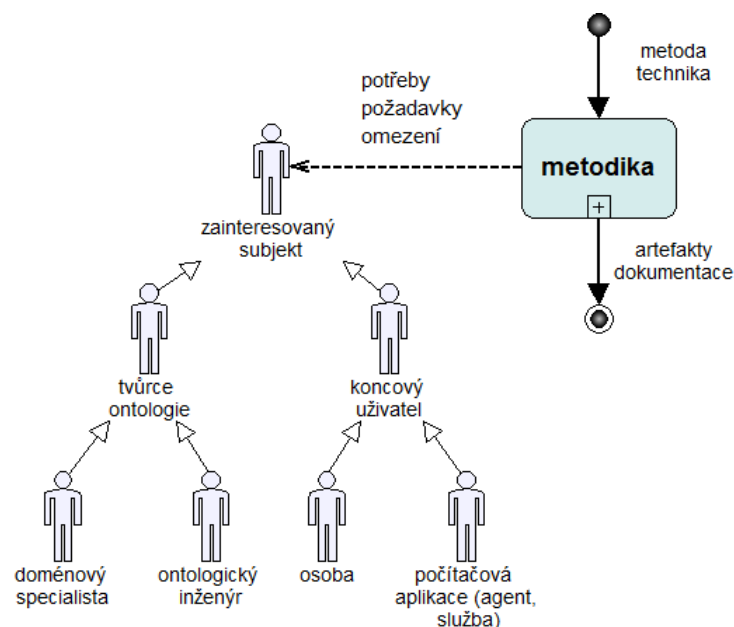
Informatické ontologie na rozdíl od těch filozofických nejsou výsledkem vědeckého zkoumání, ale jsou to artefakty, které jsou navrhovány a konstruovány. Přístup k návrhu ontologií prodělal analogický vývoj jako v případě počítačových programů a expertních systémů, jen s cca třicetiletým časovým posunem: od původního pojetí návrhu ontologie jako umění či vědy, dostupných jen několika výjimečně disponovaným jedincům, k inženýrství. Analogicky k softwarovému a znalostnímu inženýrství se zformovala nová inženýrská disciplína – ontologické inženýrství, a rovněž analogicky k CASE systémům v softwarovém inženýrství už jsou k dispozici i specializované softwarové nástroje podporující návrh ontologií (např. [Protégé](#), [Ontopia](#)). Ontologické inženýrství zahrnuje *soubor aktivit, jež se týkají procesu vývoje, životního cyklu a metod tvorby (konstrukce) ontologií, a pomůcek (např. jazyků), jež tyto aktivity podporují*. Důkazem současného stupně zralosti disciplíny jsou první publikovaná sumarizační díla – učebnice^{6, 7} a příručky^{8, 9}. Pro tuzemské zájemce o úvodní seznámení s problematikou návrhu ontologií už jsou k dispozici i dva obsáhlé přehledové texty v češtině^{10, 11}.

Stejně jako ve vědeckém výzkumu, i v inženýrském návrhu mají klíčovou roli metodiky. Z možností, jak úspěšně realizovat nějakou činnost či vyřešit problém, se jeví jako nejefektivnější právě použití osvědčené metody. Alternativa v podobě vlastního vynalézání optimálního postupu je sice šancí na kreativní, originální a potenciálně inovativní řešení, negativem jsou však vysoké náklady a ještě vyšší riziko neúspěchu. Obvyklý postup při inženýrském návrhu vypadá tedy následovně: volba metodiky – návrh, model(ování) – výroba / realizace / implementace.

Obecně lze za metodu považovat jakoukoli zaznamenanou zkušenost s prováděním nějaké činnosti či s řešením nějakého problému. Tím, že je metoda zaznamenaná, je umožněno opakované

a vícenásobné použití osvědčených postupů, nezávislé na původním nositeli. V souvislosti s metodami se v odborných textech používají v poměrně volném výkladu termíny metoda, metodika a technika a jejich významy se často protínají nebo dokonce splývají¹². Pro účely tohoto textu budeme metodiku považovat za nadřazený pojem, komplex metod a technik. Vztah metodiky a metody či techniky je vztahem více – více: v jedné metodice může být zahrnuto více metod a technik, jedna metoda či technika může být využita ve více metodikách. Technika může být věc (nástroj, pomůcka) i proces (postup), metoda je pouze proces.

Jádrem každé metodiky je tedy postup – proces, tj. zpracování vstupů na výstupy, s definovaným počátkem a koncem. Proces se může dále členit na subprocessy, což na obrázku 1 znázorňuje symbol +. Klíčovými vstupy jsou samozřejmě jednotlivé metody a techniky, které jsou v dané metodice využité. Výstupy metodiky představují především vytvořené artefakty (tj. ontologie, resp. její komponenty) a jejich dokumentace. Podstatný vliv na výslednou podobu metodiky mají potřeby, požadavky a omezení, obvykle i s ohledem na budoucí uživatele. Na obrázku 1 zastupuje jednotlivé kategorie autorů a uživatelů ontologie symbol označený jako zainteresovaný subjekt (angl. *stakeholder*). V souladu se standardem [UML](#) znázorňujeme tento vztah závislosti šipkou, mířící od závislého objektu (metodika) na objekt nezávislý (zainteresovaný subjekt). Hierarchie zainteresovaných subjektů je znázorněna symboly generalizace (dědičnosti) – trojúhelníky ukazuje od podřazeného objektu na objekt nadřazený.



Obr. 1 Procesní model metodiky

V pohledu znázorněném na obrázku 1 se tedy metodika zužuje na přehled konkrétních aktivit a jejich vzájemných vztahů. To je přístup, který uplatníme v našem přehledu metodik ontologického inženýrství, samozřejmě s vědomím toho, že samotné provedení předepsaných činností k úspěchu nestačí, a vždy bude zapotřebí brát v úvahu ještě širší kontext. Tento kontext představují jak obecné teoretické principy, paradigmaty a metody, tak konkrétní použitá pravidla, návrhové vzory, techniky, nástroje a jazykové prostředky.

			příklad
↑	teorie / vědní obor		lexikografie, logika (výroková, predikátová, algebraická, deskripční, fuzzy ad.), matematická lingvistika, systémová věda, teorie grafů, teorie informace, teorie komunikace, umělá inteligence...
	metodologie / paradigma / princip		bibliografická kontrola, ER datová analýza, fasetová analýza, konceptualizace, objektově orientovaný přístup, systémový přístup, zakotvená teorie (<i>grounded theory</i>)...
	metoda		analogie, analýza, dedukce, falzifikace, indukce, inference, klasifikace, opakované použití (<i>reuse</i>), reengineering, syntéza, verifikace ...
metodika			
↓	pravidla, návrhový vzor (<i>pattern</i>)		CCO (Cataloging cultural objects), ISO 999, ISO 5963, návrhové vzory GoF, portál Ontology Design Patterns.org, RDA (Resource description and access) ...
	technika		brainstorming, dobývání znalostí (<i>data mining</i> , <i>text mining</i>), excerpcce termínů, experiment, iterace, inkrementace, klastrování, kompetenční otázky, modelování, neuronová síť, případy užití, třídění kartiček (<i>card sorting</i>), seznam kontrolních otázek (<i>checklist</i>), sklizení (<i>harvesting</i>) zdrojů ...
	nástroj	model	diagram tříd, DTD / XSD schéma, ER diagram, rozhodovací stromy, sémantická síť, vývojový diagram ...
		software	OntoBroker, OntoLearn, Ontopia, OntoStudio, PowerDesigner, Protégé, Rational Software Architect, Visio, ...
	jazyk		CCL, CSS, CyCL, Datalog, F-logic, HTML, KIF, LISP, OCL, Ontolingua, OWL, predikátový kalkul, Prolog, RDF, RDFS, relační algebra, relační kalkul, SGML, SKOS, SPARQL, SQL, SysML, TopicMaps, UML, výrokový kalkul, WebML, XML ...

Obr. 2 Kontext metodik ontologického inženýrství

Na obrázku 2 jsou jednotlivé vrstvy kontextu metodik ontologického inženýrství řazeny za sebou od nejobecnějších po konkrétní, ve skutečnosti se ovšem prvky v nich obsažené navzájem propojují podle potřeby napříč vrstvami a bez ohledu na dané pořadí: například jazyk [OWL](#) (Web ontology language) je založen na teoretických principech deskripční logiky; metodika UP (Unified Process) využívá techniky iterace a inkrementace a návrhy podle této metodiky lze vyjádřit diagramy v jazyce UML (Unified modeling language), který vychází z objektově orientovaného paradigmatu. Příklady uvedené v tabulce jsou uvedeny v zájmu názornosti a rozhodně si nečiní nárok na úplnost. Jejich výběr byl zvolen zejména s ohledem na aplikovatelnost v ontologickém inženýrství, řazeny jsou zcela formálně podle abecedy.

V přístupu k typologii metodik ontologického inženýrství lze uplatnit poměrně značné množství pohledů, o shrnutí těch nejvýznamnějších se snažíme v přehledu v tabulce 1 a v následujícím komentáři.

Využití softwaru	<ul style="list-style-type: none"> ▪ intelektuální ▪ poloautomatizovaná ▪ automatizovaná
Postup řešení	<ul style="list-style-type: none"> ▪ top-down ▪ middle-out ▪ bottom-up
Životní cyklus	<ul style="list-style-type: none"> ▪ vodopádový ▪ iterativní
Směr vývoje	<ul style="list-style-type: none"> ▪ dopředný ▪ zpětný
Účel metodiky	<ul style="list-style-type: none"> ▪ plánovací ▪ konstrukční
Pravidla	<ul style="list-style-type: none"> ▪ rigorózní ▪ agilní
Formát	<ul style="list-style-type: none"> ▪ vědecký zákon ▪ odborný text ▪ legislativní norma ▪ technický standard
Zaměření	<ul style="list-style-type: none"> ▪ metodiky pro ontologie vytvářené „na zelené louce“ ▪ metodiky reengineeringu ontologií ▪ metodiky slučování ontologií ▪ metodiky vývoje ontologií ▪ metodiky učení ontologií ▪ metodiky hodnocení ontologií
Zdroj	<ul style="list-style-type: none"> ▪ umělá inteligence, znalostní inženýrství ▪ softwarové inženýrství ▪ knihovní a informační věda

Tab. 1 Typologie metodik ontologického inženýrství

Ontologické metodiky lze rozdělit na **intelektuální (manuální)** a **automatizované**, tj. realizované s použitím softwarových aplikací, jež jsou často vybavené nástroji umělé inteligence a umožňují např. zpracování přirozeného jazyka nebo formální konceptuální analýzu. V úvahu připadají ještě metodiky **poloautomatizované**, kombinující oba postupy.

Dalším přístupem k členění metodik je postup uplatněný při řešení problémů. Postup **top-down** směřuje od obecného celku ke konkrétním částem, **bottom-up** naopak od konkrétních částí k obecnému celku, postup **middle-out** vychází od nejdůležitějších prvků a postupuje oběma směry k nadřazeným a podřazeným úrovním. V případě návrhu ontologií se tímto způsobem specifikuje postup při tvorbě slovníku. Postup **top-down** je deduktivní – celá doména se rozčlení na několik základních kategorií a ty se postupně naplňují konkrétními pojmy, resp. termíny. Při postupu **bottom-up** se uplatňuje indukce – nejprve se shromáždí všechny termíny a ty se pak seskupují do kategorií. Postup **middle-out** vychází z pojmů považovaných za nejvýznamnější, ty se indukci zobecňují do vrcholových kategorií a dedukcí specializují na konkrétní termíny zachycující detaily dané domény.

Z hlediska životního cyklu uplatněného při návrhu je možné se setkat s různými kategorizacemi, mezi odborníky nepanuje úplná shoda ani ve významu jednotlivých termínů. Pro účely tohoto textu vycházíme z přístupu Martina Fowlera¹³, který rozlišuje dva základní typy: metodiky uplatňující vodopádový model¹⁴ (případně jeho varianty – fontánový model nebo V-model), a metodiky iterativní (a jejich varianty označované jako spirálový model¹⁵, přírůstkový neboli inkrementální model, evoluční model, prototypový model). **Vodopádový model** člení projekt podle činností (typicky na specifikaci požadavků, analýzu, návrh, testování a implementaci), jež za sebou následují v pevně stanoveném pořadí. Návraty k předchozím fázím se připouštějí, je však snaha je minimalizovat. **Iterativní model** je založen na plánovaném opakování procesů. Vychází se z hypotézy, že opakovaným řešením (verzováním, prototypováním) lze dospět ke kvalitnějšímu

výsledku. Zpravidla se před zahájením iterací navrhne základní architektura systému, a poté se přírůstkově řeší jednotlivé části (komponenty, funkcionality). Výstupem každé iterace má být nový přírůstek rozsahu a kvality projektu v podobě použitelné části navrhovaného systému. V průběhu dalších iterací se nejen řeší další přírůstek, ale znovu se revidují už dokončené části.

Nejobvyklejším způsobem inženýrského návrhu je postup směrem *top-down* od modelu k implementaci – tzv. **dopředné inženýrství** (*forward engineering*). Pokud se z hotové funkční aplikace ve směru *bottom-up* dodatečně odvozuje abstraktní konceptuální model, hovoří se o tzv. **zpětném inženýrství** (*reverse engineering*). Oba směry vývoje mohou být v současné době podporovány softwarovými prostředky a do jisté míry automatizovány.

Podle zaměření na konkrétní aktivity v průběhu projektu se metodiky člení na **plánovací** (výsledkem jsou opakovatelné procesy a postupy) a **konstrukční** (výsledkem jsou znovupoužitelné modely, resp. pravidla pro jejich konstruování).

V souvislosti s množstvím a závazností pravidel obsažených v metodice se mluví o metodikách **rigorózních** (striktních), kladoucích důraz na dobře definované procesy vývoje, a o **agilních** (*agile, light, lightweight*) metodikách, jež za nejdůležitější faktor úspěchu projektu považují kvalitu zúčastněných osob a jejich efektivní spolupráci, a regulaci procesů se snaží minimalizovat ve prospěch kreativních ad hoc řešení a pružných reakcí na změny v průběhu projektu.

Kromě obsahu je možné uvažovat i formáty, v nichž jsou metodiky zaznamenány. Ty mohou mít rozpětí od **vědeckého zákona** přes **odborné texty** až po pravidla a standardy na různé úrovni závaznosti. V úvahu připadají jak dobrovolné **technické normy** – např. ISO nebo ČSN, tak **legislativní normy** – např. zákony jednotlivých zemí či direktivy EU. Podle rozsahu působnosti je možné se setkat s pravidly a standardy na mezinárodní, národní, odvětvové úrovni (např. vydávané organizacemi [IFLA](#), [CIDOC](#), [IEEE](#)), s podnikovými předpisy a s tzv. *best practices*.

Metodiky ontologického inženýrství už se stačily typologicky rozvrstvit podle toho, které specifické úlohy při vývoji ontologie plní. Kromě kompletního vývoje celé ontologie „od nuly“, „**na zelené louce**“ (*from scratch*) jsou k dispozici i metodiky specializované na úlohy **reengineeringu** (zásadní přepracování existující ontologie), metodiky **slučování ontologií** a metodiky **vývoje** či **učení** existujících ontologií (*ontology evolution, ontology learning*), jež se věnují postupnému doplňování ontologie o další pojmy a vztahy, často realizovanému s podporou specializovaného softwaru. Specifickou kategorií představují také metodiky pro **hodnocení ontologií**.

Již bylo konstatováno, že ontologie navazují na předchozí aktivity v oblasti organizace a zpracování znalostí. Rovněž metodiky ontologického inženýrství využívají osvědčené metody příbuzných oborů, nejčastěji umělé inteligence a znalostního inženýrství, softwarového inženýrství a knihovní a informační vědy.

Umělá inteligence a znalostní inženýrství jsou obory zaměřené na návrh znalostních a expertních systémů, případně znalostníchází. Jejich klíčovými metodikami jsou logická pravidla (*rules*) – procedurální reprezentace znalostí produkčními pravidly typu předpoklad – závěr (*if – then*), a rámce (*frames*)¹⁶ – systém reprezentace objektů v bázi znalostí prostřednictvím položek (*slots*), sloužících k popisu jejich vlastností. Další metodou ověřenou v této oblasti a uplatnitelnou při návrhu ontologií je formální konceptuální analýza (*formal concept analysis*)¹⁷ – metodika automatické derivace ontologií z datových objektů a jejich vlastností. Značný význam pro ontologické inženýrství mají vyspělé metody získávání znalostí (*knowledge acquisition*)¹⁸, jež pomáhají v etapě, která je obecně považována při návrhu ontologie za kritickou. Patří mezi ně techniky získávání (elicitace) znalostí od expertů (interview, *brainstorming*, třídění kartiček, repertoárové tabulky) a automatizované nebo poloautomatizované získávání znalostí z dat (např. *data mining*).

Softwarové inženýrství je disciplína zaměřená na návrh softwaru, případně informačního systému (systémové inženýrství, systémová integrace). Typickým produktem je softwarová aplikace nebo databáze. Soustředí se zpravidla na unikátní aplikace, navrhované na míru specifickým konkrétním podmínkám a uživatelským potřebám (tzv. *usability* – individuální použitelnost). Softwarové

inženýrství má k dispozici propracované metodiky datové analýzy (např. ERA modelování, diagramy tříd) a metodiky životního cyklu softwarového projektu – např. MMDIS, [RUP](#) – Rational Unified Process, Unified Process, Scrum, Crystal, Extrémní programování.

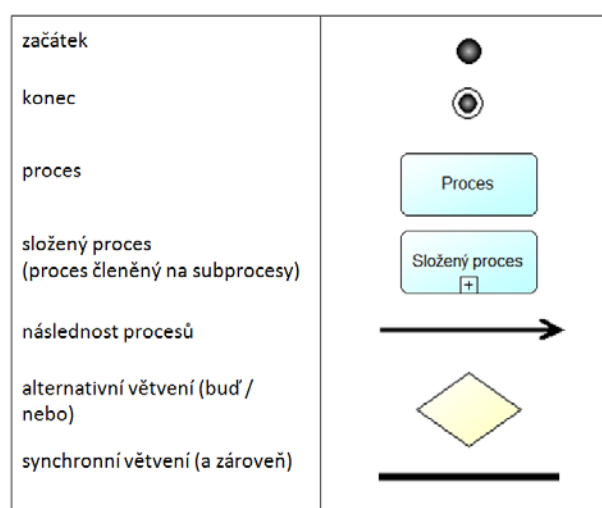
Knihovní věda a informační věda uplatňují své poznatky v aplikační oblasti návrhu a implementace knihovních a informačních systémů. Produkty významné z hlediska ontologického inženýrství jsou tradičně označovány jako selekční jazyky, v současné době se pro ně často používá termín systém organizace znalostí (*KOS – knowledge organization system*). Knihovnické ontologie (MDT, tezaury, [CIDOC CRM](#)) zůstávají zatím s výjimkou medicíny poněkud stranou pozornosti odborníků na ontologie. Jejich předností nebývá propracovaný systém vztahů, ale bohatost lexikálního obsahu, který zase mnohdy chybí „pravým“ ontologiím vzniklým v laboratorních podmínkách. Soustředí se zpravidla na dosažení širokého konsensu, jejich ambicí je tedy na rozdíl od softwarového inženýrství opakovatelná použitelnost (*reusability*). K tomuto účelu mají často vypracovaný velmi důkladný a obsáhlý soubor pravidel, uplatňovaných např. při tvorbě slovníku. Metodiky využitelné při tvorbě ontologií jsou jak kvantitativní (bibliometrie, citační analýza), tak kvalitativní, založené na analýze obsahu. Významným přínosem informační vědy je fasetová analyticko-syntetická metoda¹⁹, v praxi osvědčené metody automatické obsahové analýzy a indexování²⁰ a četné metodiky tvorby tezauru (viz část 3.6). V neposlední řadě přispěla knihovní věda k propracování principu označovaného jako tzv. *warrant* – přednost pragmatického aspektu před teoretickými principy; je-li to pro praktické účely potřeba, princip *warrant* umožňuje „ospravedlnit“ i vytvoření ontologie toho, co „není“ – např. pohádkových bytostí, o nichž se píše v literatuře.

3. Přehled metodik návrhu ontologií

Všechny metodiky uvedené v našem přehledu spadají do kategorie metodik určených pro návrh doménových ontologií, uplatňují dopředné inženýrství a jsou určeny pro návrh nových ontologií typu „na zelené louce“. V zásadě lze každou z nich označit za rigorózní, míra rigoróznosti se však v jednotlivých metodikách liší. Krajní pól představuje metodika Uschold – King / Enterprise, jež nabízí pouze základní kostru určenou k naplnění individuálně zvolenými postupy. Na opačné straně spektra stojí např. metodika Methontology a svým způsobem i metodiky tvorby tezauru s detailně vypracovanými pravidly pro řešení jednotlivých jazykových problémů. Podle své příbuznosti s vědními disciplínami jsou metodiky v přehledu rozděleny do tří skupin. V první skupině jsou uvedeny metodiky, jež mají svůj aparát založen na metodách umělé inteligence a znalostního inženýrství: Dnes již klasické a vzájemně se doplňující metodiky Grüninger – Fox / TOVE a Uschold – King / Enterprise, obě vycházející z prostředí podnikové ekonomiky. Doplňuje je obecně zaměřená metodika Noy – McGuinness, která na rozdíl od předchozích dvou metodik, založených na logických pravidlech, vychází z konceptu založeného na rámcích. Jako příklady metodik založených na principech softwarového inženýrství jsou představeny metodiky Methontology a UPON. Přehled uzavírají dvě metodiky pocházející z prostředí knihovních a informačních systémů: metodiky tvorby tezaurů, jež mají dlouhou tradici (datují se už od 70. let 20. století) a jako jediné mají statut mezinárodně přijatých ISO norem, a metodika Denton – Spiteri, postavená na principech fasetové analýzy, jejíž kořeny v podání S. R. Ranganathana sahají až do 20. let 20. století.

I když je náš přehled zaměřen informativně a jeho cílem není srovnávat ani hodnotit jednotlivé metodiky, snažili jsme se uplatnit jednotnou strukturu popisu metodik, která by měla umožnit případnému zájemci jejich srovnání. U každé metodiky je uveden její producent nebo autor (autoři) a odkazy na zdrojové dokumenty, rok vzniku či publikování a stručná charakteristika určení a specifík metodiky. Jádrem popisu představuje procesní model metodiky, který tvoří slovní popis a procesní graf (příp. grafy) klíčových fází. V závěrečném tabulkovém přehledu (tabulka 3) je uvedeno zařazení popisovaných metodik podle typologie uvedené v části 2.

Na obrázku 3 jsou uvedeny grafické konvence, použité v procesních grafech:



Obr. 3 Použité grafické symboly

Názvy procesů se v jednotlivých metodikách liší – používají se označení etapa, fáze, krok, činnost, aktivita, pracovní postup (*workflow*), úloha, stupeň (*stage*) ad. V našem přehledu používáme bez ohledu na původní terminologii autorů metodik jednotně termín „fáze“. Obdobný terminologický problém nastává u pojmenování základního konstrukčního prvku ontologie v jednotlivých metodikách – na konceptuální úrovni se označuje jako znalost / poznatek, pojem, třída, entita, objekt, instance, na úrovni formální (znakové) se lze setkat s označením lexikální jednotka / lexém, slovo, termín, deskriptor. I v tomto ohledu se snažíme o sjednocení používané terminologie ve prospěch označení „pojem“ a „termín“, ostatní termíny zachováváme v popisech metodik jen tehdy, pokud je to nezbytně nutné pro pochopení smyslu.

Poznámka k názvům metodik: U metodik ontologického inženýrství je poměrně běžné, že jim jejich autoři nedávají individuální názvy. Z našeho souboru jsou známy pod oficiálními názvy pouze metodiky Methontology a UPON. Pro ostatní metodiky používáme názvy odvozené ze jmen jejich autorů, u označení metodiky tvorby tezauru jsme zůstali u popisného názvu.

3.1 Grüninger – Fox / TOVE

Autoři: Michael Grüninger a Mark S. Fox (Enterprise Integration Laboratory, University of Toronto).

Rok vzniku / publikování: 1995.

Určení, specifika: Metodika vychází ze zkušeností autorů s projekty zaměřenými na návrh znalostního informačního systému o podniku, založenými na principech procesního managementu (Common Sense Enterprise Model, Enterprise Engineering, TOVE – Toronto Virtual Enterprise). Navrhuje přístup ke konstrukci ontologie založený na pravidlech predikátové logiky prvního řádu. V životním cyklu se uplatňuje vodopádový model s prvky iterace, autoři doporučují prototypový postup ve směru *middle-out*.

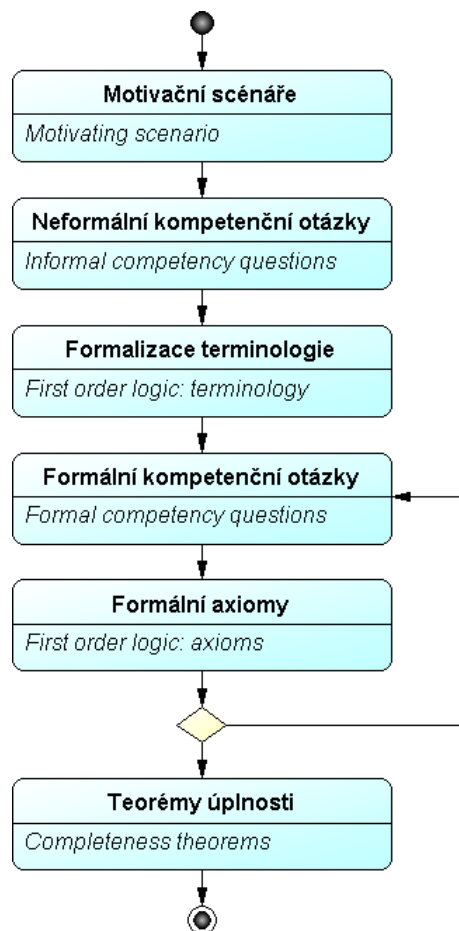
Klíčové fáze:

- Motivační scénáře: Obvykle narativní formou příběhu (*story*) zachycené problémy, jimž se má ontologie věnovat, včetně intuitivních návrhů na jejich řešení.
- Neformální kompetenční otázky: Pro každý objekt v ontologii má existovat otázka, na niž má odpovídat. Autoři doporučují kompetenční otázky stratifikovat, tak aby otázky na vyšší úrovni bylo možné zodpovědět s použitím výsledků jednodušších otázek na nižší úrovni.
- Formalizace terminologie: Termíny navržené k zařazení do ontologie mají být specifikovány s použitím predikátové logiky prvního řádu, případně v jazyce [KIF](#) (Knowledge interchange

format). V rámci vymezeného univerza diskursu se definují objekty, které jsou reprezentovány prostřednictvím konstant a proměnných. Pro objekty se definují atributy jakožto unární predikáty a vztahy mezi nimi se definují pomocí n-árních (např. binárních) predikátů.

- Formální kompetenční otázky: S pomocí formalizované terminologie se neformální kompetenční otázky převedou na formální.
- Formální axiomy: Definují se formální axiomy, jež s použitím predikátové logiky prvního řádu specifikují definice termínů v ontologii a omezení týkající se jejich interpretace. Axiomy se ověřují prostřednictvím formálních kompetenčních otázek, tato fáze je tedy iterativní.
- Teorémy úplnosti: S využitím formálních axiomů se definují podmínky, za nichž je možné prohlásit řešení kompetenčních otázek za kompletní a tím prověřit úplnost ontologie.

Procesní graf (obrázek 4):



Obr. 4 Grüninger – Fox

Zdroj:

GRÜNINGER, Michael a FOX, Mark S. Methodology for the design and evaluation of ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95* [online], Montreal, 1995 [cit. 19.4.2011]. 10 s. Dostupné z: <http://www.eil.utoronto.ca/enterprise-modelling/entmethod/index.html>

3.2 Uschold – King / Enterprise

Autoři: Michael Uschold a Martin King (Artificial Intelligence Applications Institute, University of Edinburgh).

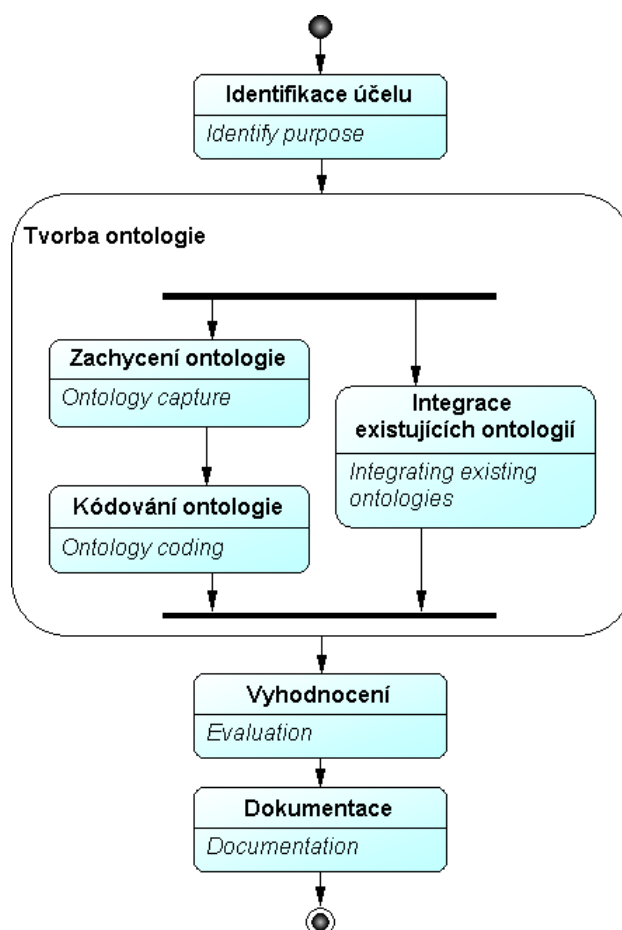
Rok vzniku / publikování: 1995.

Určení, specifika: Metodika je založená na zkušenostech autorů z realizace ontologie Enterprise pro oblast podnikové ekonomiky a doplněná o náměty a poznatky získané studiem dosavadní odborné literatury (do r. 1995). Autoři nabízejí základní skelet metodiky v podobě několika fází, určený k dalšímu konkretizování a k dopracování podle potřeb konkrétních aplikačních oblastí. Ilustrují využití navržené skeletové metodiky případovou studií popisující detailně aktivity realizované v průběhu návrhu ontologie. V životním cyklu se přiklání k postupu *middle-out*.

Klíčové fáze:

- Identifikace účelu: Definování zamýšleného použití ontologie a jejích uživatelů, včetně způsobu využití ontologie uživatelem (např. typ softwaru, který bude umožňovat přístup k ontologii). K vymezení věcného zaměření autoři doporučují zformulovat kompetenční otázky.
- Tvorba ontologie: Zahrnuje tři fáze – zachycení a kódování ontologie, a souběžně realizovanou fázi integrace existujících ontologií.
 - Zachycení ontologie: Identifikace klíčových pojmů a vztahů v zájmové oblasti. Tvorba přesných a jednoznačných textových definic těchto pojmů. Nalezení a odsouhlasení termínů označujících definované pojmy a vztahy. Relevantní oblasti, z níž je doporučeno v této etapě čerpat, jsou metody získávání znalostí, známé z umělé inteligence.
 - Kódování ontologie: Vyjádření zachycené konceptualizace v nějakém formálním jazyce.
 - Integrace existujících ontologií: V průběhu zachycování i kódování ontologie se zvažuje, zda je vhodné a možné použít již existující ontologie.
- Vyhodnocení: Posouzení, zda navržená ontologie odpovídá specifikovaným požadavkům, kompetenčním otázkám a/nebo realitě.
- Dokumentace: Autoři konstatují význam dokumentace pro použitelnost ontologie a navrhují rozpracovat pravidla pro její tvorbu, přizpůsobená různým typům ontologií.

Procesní graf (obrázek 5):



Obr. 5 Uschold – King

Zdroj:

USCHOLD, Michael a KING, Martin. Towards a methodology for building ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95, Montreal, 1995* [online]. Edinburgh: Artificial intelligence applications institute, University of Edinburgh, 1995 [cit. 19.4.2011]. 13 s. AIAI-TR-183. Dostupné z: <http://en.scientificcommons.org/43164996>

3.3 Noy – McGuinness

Autorky: Natalya F. Noy a Deborah L. McGuinness (Stanford Center for Biomedical Informatics Research, Stanford University School of Medicine, Stanford).

Rok vzniku / publikování: 2001.

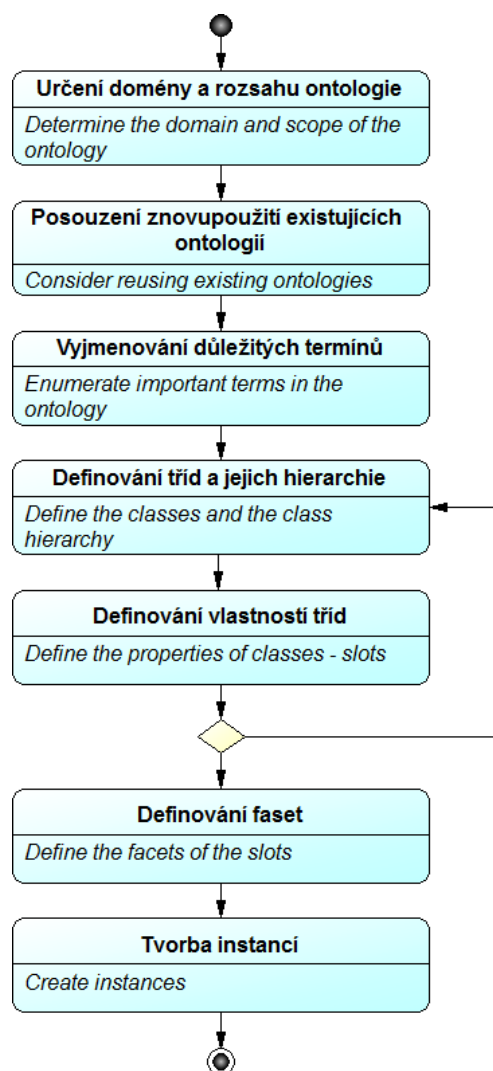
Určení, specifika: Metodika vypracovaná s využitím zkušeností autorek s navrhováním ontologií v prostředí softwarových platforem Protégé, [Ontolingua](#) a [Chimaera](#). Uplatňují přístup ke konstrukci ontologie založený na rámcích, interpretovaný tak, aby byl vhodný i pro začátečníky. Nabízí vodopádový postup s prvky iterace.

Klíčové fáze:

- Určení domény a rozsahu ontologie: Definování domény, oblastí použití a uživatelů ontologie. Zformulování kompetenčních otázek.
- Posouzení znovupoužití existujících ontologií: Autorky konstatují, že pro většinu domén už existují ontologie uložené ve veřejně dostupných knihovnách ontologií, jež lze importovat, rozšířit či modifikovat, a doporučují to provést.

- Vyjmenování důležitých termínů: Cílem je vytvořit co nejobsáhlejší seznam termínů, které budou obsaženy ve výrocích nebo ve vysvětleních poskytovaných ontologií uživateli, zatím bez řešení jejich sémantických charakteristik a vztahů.
- Definování tříd a jejich hierarchie: Podle zaměření návrháře, eventuálně podle specifik dané domény se použije buď postup *top-down*, *bottom-up* nebo *middle-out*. Ze seznamu termínů se vyberou termíny označující objekty se samostatnou existencí a označí se jako třídy. Mezi třídami se definují generické hierarchické vztahy s uplatněním dědičnosti. Obvykle se tato fáze těsně propojuje s fází následující a realizuje se iterativně. Typický průběh vypadá tak, že se vybere několik kandidátních tříd a pro ně se určí jak hierarchické vztahy, tak jejich vlastnosti. Poté se pokračuje další skupinou tříd.
- Definování vlastností tříd: Z termínů, které nebyly označeny jako třídy, se vyberou ty, jež představují vlastnosti některé z definovaných tříd. Rozlišují se typy vlastností: přirozené / vnitřní (např. barva), vnější (např. název), části strukturovaných objektů a vztahy k jiným objektům.
- Definování faset: Fasetami se rozumí v této metodice „vlastnosti vlastností“, popisující datový typ, povolené hodnoty, počet hodnot (kardinalitu) a případné další parametry. U vlastností, které vyjadřují vztahy, se určí definiční obor (*domain*) a obor hodnot (*range*).
- Tvorba instancí: Vytvoření individuálních instancí tříd a naplnění jejich vlastností hodnotami.

Procesní graf (obrázek 6):



Obr. 6 Noy – McGuinness

Zdroj:

NOY, Natalya F. a McGUINNESS, Deborah L. *Ontology development 101: A guide to creating your first ontology* [online]. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001 [cit. 19.4.2011]. 25 s. Dostupné z: http://protege.stanford.edu/publications/ontology_development/ontology101.html a z: <http://www-ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html>

3.4 Methontology

Producent: [Ontology Engineering Group](#) (Universidad Politécnica de Madrid)

Rok vzniku / publikování: 1997 (popis metodiky je zpracován podle verze publikované v roce 2004).

Určení, specifika: Východiskem metodiky je standard IEEE 1074 pro návrh životního cyklu softwarových procesů v dnes již neplatné verzi z roku 1996²¹. V souladu s tímto standardem jsou aktivity životního cyklu návrhu ontologie rozčleněny na tři paralelně probíhající skupiny procesů: správa, vývoj, podpora. Pro fáze vývoje je navrženo pořadí provádění, u fází věnovaných správě a podpoře ontologie se předpokládá, že jsou realizovány průběžně během celého životního cyklu projektu. Životní cyklus je založen na prototypování, počítá se s iterativním postupem ve směru *middle-out*. Na podporu metodiky byly vyvinuty softwarové editory ODE a WebODE, je použitelná i s jinými nástroji (např. Protégé). Je doporučena organizací [FIPA](#) (Foundation for Intelligent Physical Agents).

Klíčové fáze:

Aktivity řízení a správy ontologie – plánovací fáze se uskutečňuje na počátku každého vývojového cyklu, kontrola a zajištění kvality probíhají průběžně.

- Plánování: Identifikace úloh, jež mají být vykonány, jejich pořadí a potřebného času a zdrojů.
- Kontrola: Ověření, zda byly naplánované úlohy provedeny.
- Zajištění kvality: Kontrola kvality výstupů (ontologie, software, dokumentace).

Aktivity orientované na vývoj ontologie – člení se na předvývojovou fázi, vlastní vývoj a post-vývojovou fázi.

- Předvývojová fáze: V jejím průběhu se vypracují studie prostředí a studie proveditelnosti.
 - Studie prostředí: Určení platformy, na níž bude ontologie provozována, a aplikací, s nimiž má být integrována.
 - Studie proveditelnosti: Posouzení vhodnosti a uskutečnitelnosti záměru vytvořit ontologii.
- Vývoj: Fáze vývoje se dále člení na iterativně prováděné fáze specifikace požadavků, konceptualizace, formalizace a implementace.
 - Specifikace: Určení účelu, uživatelů a zamýšleného použití ontologie.
 - Konceptualizace: Vytvoření konceptuálního modelu z materiálu získaného v průběhu fáze získávání znalostí. Nejvýznamnější fáze tvorby ontologie – autoři doporučují v jejím průběhu vypracovat následující komponenty:
 - Výkladový slovník termínů: Zpracování definic a seznamu synonym a zkratk pro všechny termíny, jež budou obsaženy v ontologii.
 - Pojmové taxonomie: Výstupem je jedna nebo více taxonomií zachycující hierarchické vztahy mezi pojmy.
 - Diagramy ad hoc binárních vztahů: Vytvoří se diagramy zachycující ad hoc binární vztahy

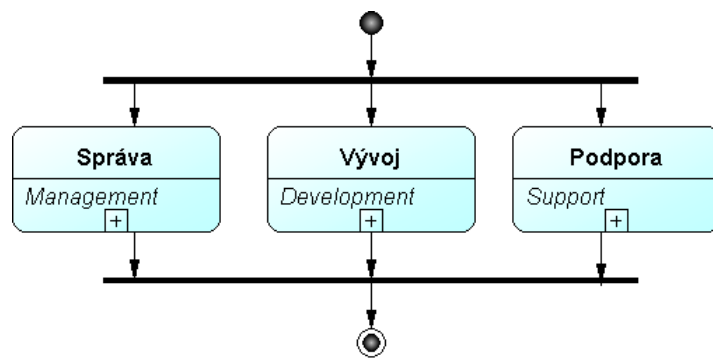
mezi pojmy v ontologii navzájem a mezi pojmy v ontologii a pojmy z jiných ontologií.

- Výkladový slovník pojmů: Slovník pojmů zahrnuje jejich instance (nepovinné), atributy tříd a instancí a ad hoc vztahy.
- Popis ad hoc binárních vztahů: Detailní popis každého vztahu, vyskytujícího se v diagramu.
- Popis atributů instancí: Sestavení tabulky s podrobným popisem všech atributů každé instance.
- Popis atributů tříd: Sestavení tabulky s podrobným popisem všech atributů každé třídy.
- Popis konstant: Podrobný popis všech konstant, jež specifikují informace týkající se znalostní domény, a jež mohou být používány ve vzorcích.
- Formální axiomy: Identifikace a přesný popis formálních axiomů vyskytujících se v ontologii. Autoři doporučují zpracovat jejich formulaci prostřednictvím predikátové logiky prvního řádu.
- Pravidla: Identifikace a přesný popis pravidel vyskytujících se v ontologii. Autoři doporučují pro zápis pravidel konstrukci *if – then*.
- Instance: Po dokončení konceptuálního modelu se mohou definovat relevantní instance. Nepovinná fáze.
- Formalizace: Vyjádření konceptuálního modelu v některém formálním jazyce. Pokud je při konceptualizaci používán softwarový nástroj typu WebODE, lze tuto fázi provést automatizovaně. Nepovinná fáze.
- Implementace: Vytvoření počítačem zpracovatelného modelu ve zvoleném ontologickém jazyce.
- Postvývojová fáze: Zahrnuje údržbu a případné (znovu)použití.
 - Údržba: Aktualizace a opravy ontologie podle potřeby.
 - (Znovu)použití: Využití existující ontologie jinými ontologiemi nebo aplikacemi.

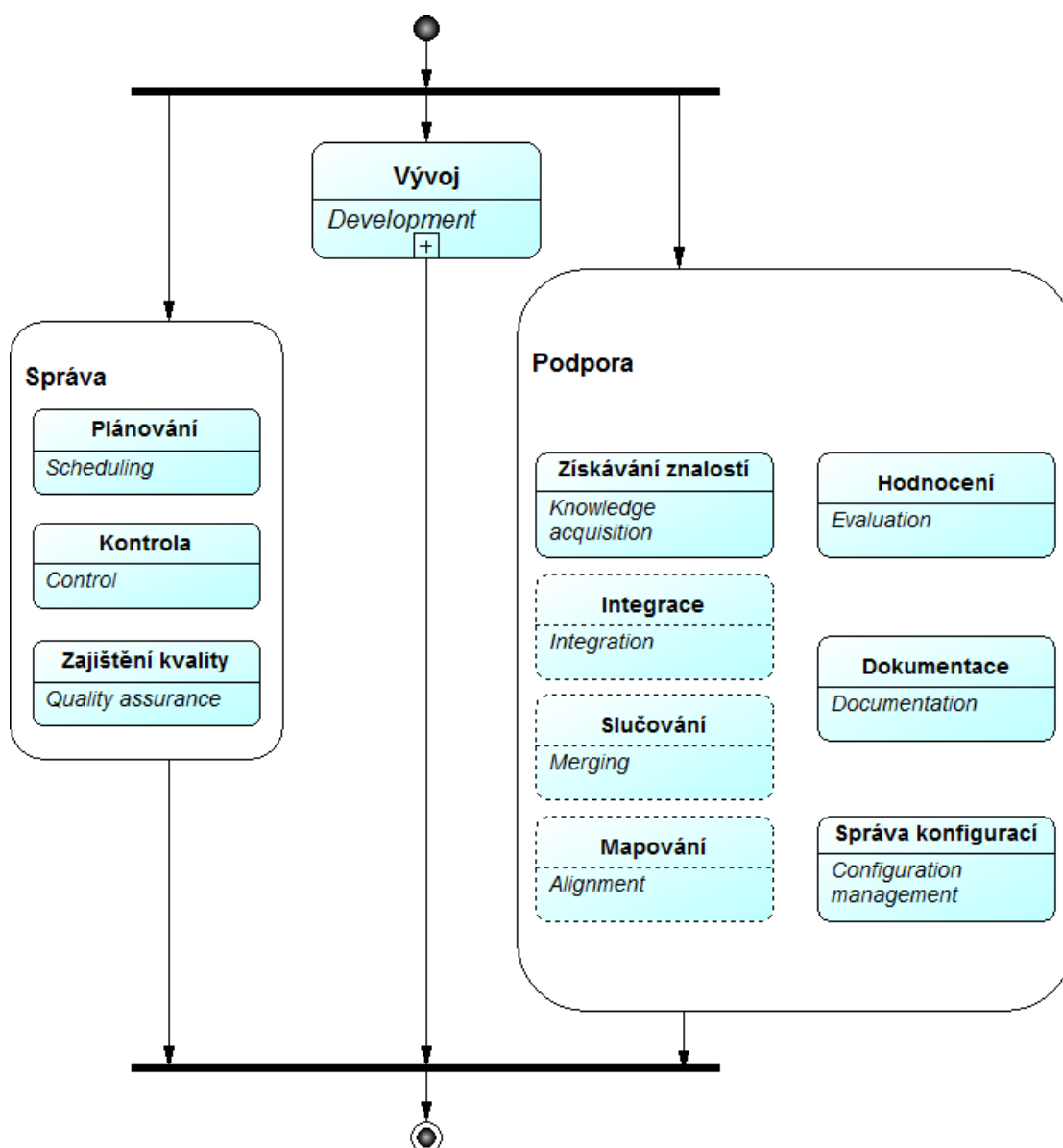
Aktivita podpory ontologie – veškeré aktivity se uskutečňují podle potřeby souběžně s vývojovými fázemi ontologie.

- Získávání znalostí: Získání znalostí od doménových expertů nebo nějakým automatizovaným procesem (učení ontologií).
- Hodnocení: Technické posouzení ontologie, jejího softwarového prostředí a dokumentace.
- Integrace: Realizuje se v případě, že se vytváří nová ontologie s využitím již existujících ontologií.
- Slučování: Realizuje se v případě, že se vytváří nová ontologie sloučením již existujících ontologií nebo jejich částí, přičemž dochází k sémantickému sjednocení.
- Mapování: Realizuje se v případě, že jsou zdrojové ontologie ponechány beze změny a mezi odpovídajícími částmi jsou definovány vazby.
- Dokumentace: Úplný popis každé dokončené fáze a vytvořeného produktu.
- Správa konfigurací: Záznam všech vytvořených verzí dokumentace i ontologie samotné.

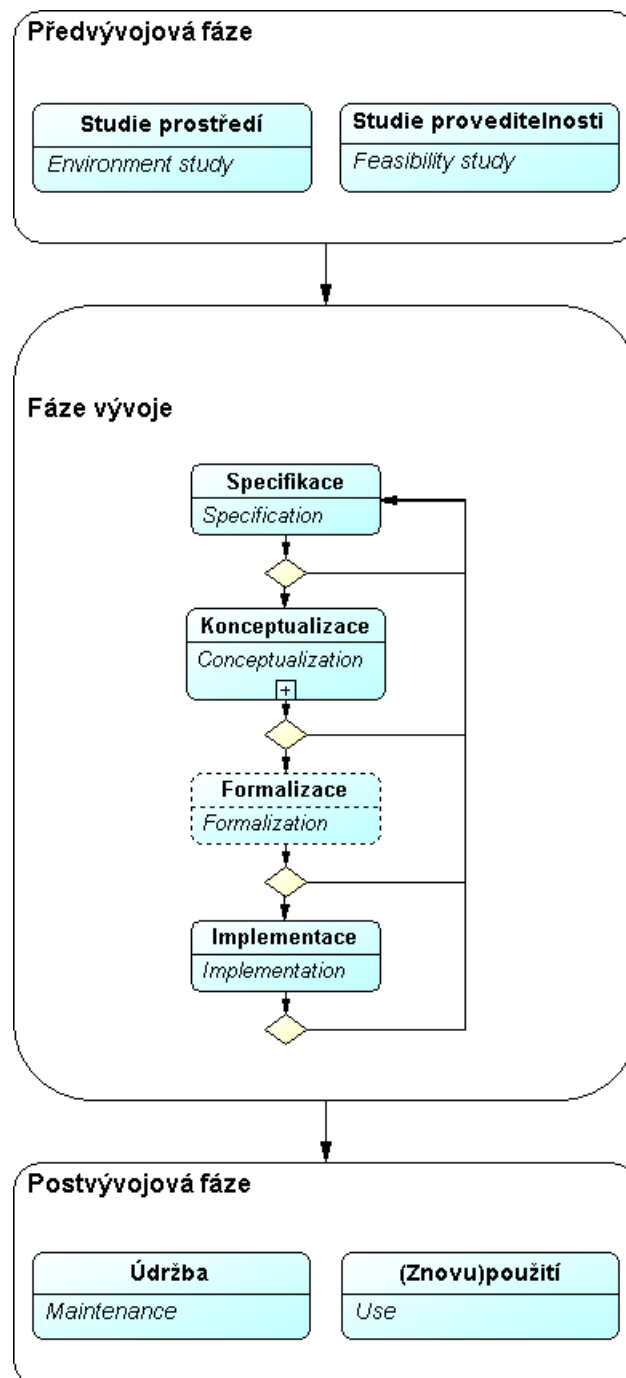
Procesní grafy (obrázky 7 – 10):



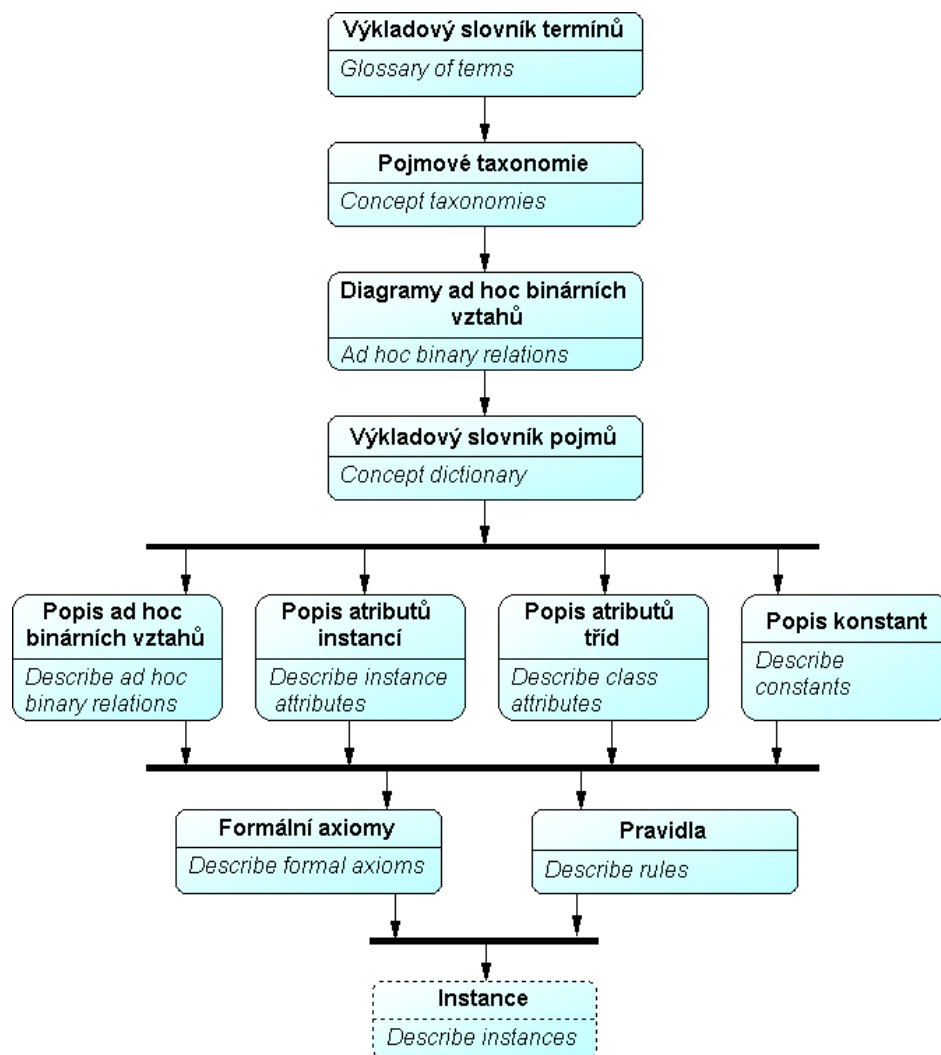
Obr. 7 Obecný rámec metodiky Methontology



Obr. 8 Methontology – detailní členění fází Správa a Podpora



Obr. 9 Methontology – detailní členění fáze Vývoj



Obr. 10 Methontology – detailní členění fáze Konceptualizace

Zdroje:

FERNÁNDEZ-LÓPEZ, Mariano, GÓMEZ-PÉREZ, Asunción a JURISTO, Natalia. METHONTOLOGY: From ontological art towards ontological engineering. In *Spring Symposium on Ontological Engineering of AAAI*, 24.-26.3.1997 [online]. Stanford University, California, 1997 [cit. 24.4.2011], s. 33-40. AAAI Technical Report SS-97-06. Dostupné z: <http://www.aaai.org/Papers/Symposia/Spring/1997/SS-97-06/SS97-06-005.pdf>

GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano a CORCHO, Oscar. *Ontological Engineering: With examples from the areas of knowledge management, e-commerce and the Semantic Web*. 1. ed. London: Springer, 2004. Kapitola 3.3.5, METHONTOLOGY, s. 125–142.

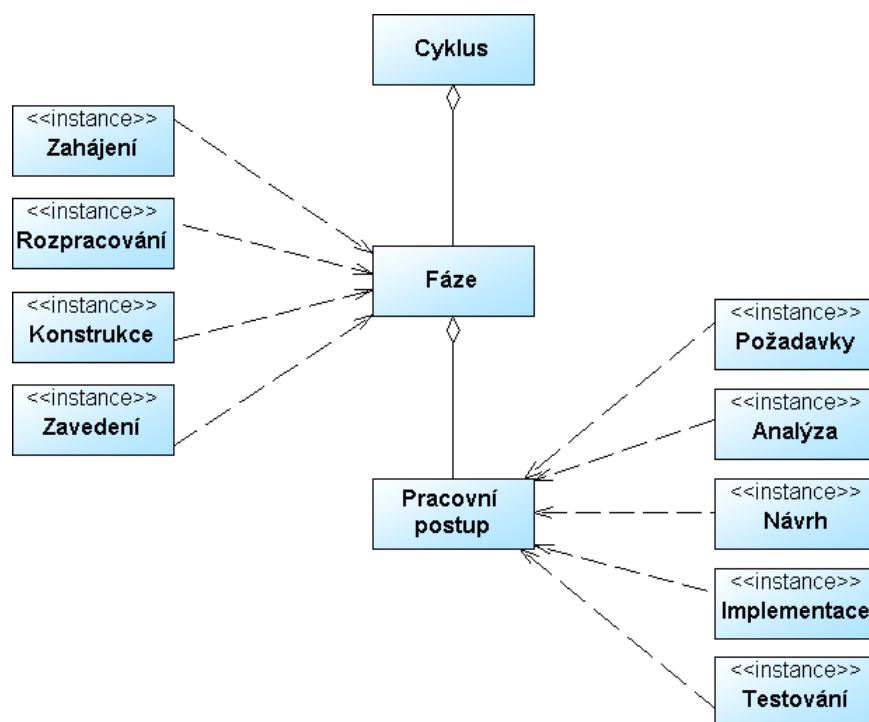
3.5 UPON: Unified Process for ONtology

Autoři: Antonio De Nicola a Michele Missikoff (Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche, Roma), Roberto Navigli (Dipartimento di Informatica, Università di Roma „LaSapienza“, Roma).

Rok vzniku / publikování: 2005 (popisována je verze publikovaná v roce 2009).

Určení, specifika: Metodika UPON byla vyvinutá v rámci projektů Interop NoE a Athena IP, jež jsou součástí 6. rámcového programu EU. Jejím zdrojem je metodika Unified Process (UP)²², která je v současnosti považována za standard ve vývoji rozsáhlých softwarových aplikací.

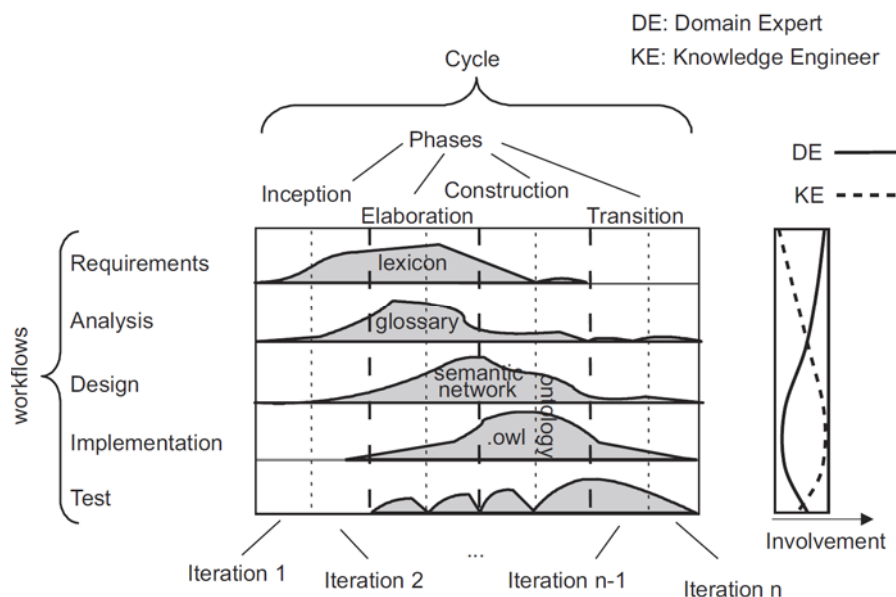
Charakteristickými rysy metodiky UP jsou: orientace na uživatele (vývoj systému řízený případy užití), iterativní a inkrementální vývoj, prototypování, znovupoužití (*reuse*) existujících návrhů, vzorů i komponent. Nabízí relativně komplikovaný životní cyklus, který ovšem právě díky své komplexnosti lépe odpovídá reálnému průběhu softwarových projektů, než je tomu například u jednoduchého vodopádového postupu. Strukturu softwarového projektu tvoří cykly, členěné na čtyři fáze (*inception* – zahájení, *elaboration* – rozpracování, *construction* – konstrukce, *transition* – zavedení), a každá fáze se vnitřně člení na stejných pět pracovních postupů (*requirements* – požadavky, *analysis* – analýza, *design* – návrh, *implementation* – implementace, *testing* – testování). Jak cykly, tak fáze mohou být prováděny opakovaně (iterativně). Iterace na úrovni celého cyklu vede k vývoji nové verze. Grafickým vyjádřením této struktury je schéma na obrázku 11. V souladu se standardem UML znázorňujeme v diagramu vztah celek – část šipkou ve tvaru kosočtverce, mířící od části k celku. Konkrétní fáze a pracovní postupy jsou připojeny vztahem závislosti znázorněným šipkou vedoucí od instance k její třídě.



Obr. 11 Struktura metodiky Unified Process

Metodika UPON přebírá beze změny strukturu procesu UP včetně názvů fází a pracovních postupů, které však naplňuje konkrétním obsahem přizpůsobeným pro účely návrhu ontologie (viz obr. 12). Je určena pro návrh rozsáhlých i menších doménových ontologií, určených pro oblast přesně vymezenou konkrétními případy užití. Při návrhu ontologie se počítá s dvěma klíčovými typy odborníků – znalostními (ontologickými) inženýry a doménovými specialisty, jimž jsou v procesu vývoje určeny aktivity, za něž zodpovídají. Přírůstkový charakter metodiky se projevuje v postupném vytváření a doplňování výstupů jednotlivých fází: 1) fáze zahájení: slovník (*lexicon* = seznam termínů, 2) fáze rozpracování: výkladový slovník (*glossary* = termíny + definice), 3) fáze

konstrukce: sémantická síť (*semantic network* = slovník + vztahy), 4) fáze zavedení: doménová ontologie. Tyto opakující se revize a zdokonalování jednotlivých výstupů jsou vyznačeny v iteracích na obrázcích 12 a 13.



Obr. 12 Obecný rámec metodiky UPON²³

Klíčové fáze:

Fáze Požadavky – cílem je specifikace sémantických potřeb uživatelů a uživatelského pohledu na znalosti, jež mají být zachyceny v ontologii. Vstupem jsou rozhovory s doménovými experty a dokumenty týkající se aplikační oblasti ontologie. Výstupem jsou kompetenční otázky, případy užití a aplikační slovník.

- Určení zájmové domény a rozsahu ontologie: Vymezení té části reality, na niž se ontologie zaměří. Pokud je doména rozsáhlá, určí se subdomény. Nástrojem jsou pojmy a ontologické závazky²⁴.
- Definování účelu nebo motivačního scénáře s uživateli a jejich záměry: Vymezení uživatelů ontologie, jejich typologie a úloh, jež bude ontologie pro uživatele řešit.
- Vypracování kresleného scénáře: Načrtnutí pořadí aktivit, jež se odehrávají v rámci konkrétního scénáře.
- Vytvoření aplikačního slovníku: Shromáždění terminologie od doménových expertů a z dokumentů používaných v aplikační doméně. Možnost využití kreslených scénářů a nástrojů pro extrakci znalostí z textových dokumentů (např. [OntoLearn](#)).
- Identifikace kompetenčních otázek: Formulace otázek na pojmové úrovni, jež má být ontologie schopna zodpovídat. Hlavními typy jsou otázky směřující k vyhledávání (informačních) zdrojů a otázky, jejichž cílem je dosažení sémantické dohody o významu zdrojů.
- Modelování případů užití: Určení funkcionalit, jež má ontologie poskytovat, aby umožnila zodpovězení kompetenčních otázek. Autoři doporučují použít modely případů užití, definované ve standardním modelovacím jazyce UML.

Fáze Analýza – cílem je detailní rozpracování a strukturování požadavků. Vstupem je aplikační slovník a existující externí zdroje (dokumenty, standardy, ontologie). Výstupem jsou diagramy tříd, diagramy aktivit a referenční výkladový slovník.

- Získání doménových zdrojů a tvorba doménového slovníku: Doménový slovník tvoří terminologie používaná v zájmové doméně, převážně získaná analýzou existujících externích

dokumentových zdrojů, jako jsou zprávy, manuály, standardy, slovníky, tezaury a dostupné ontologie. Tvorba doménového i aplikačního slovníku může být podpořena automatizovanými nástroji pro *text mining*.

- Vytvoření referenčního slovníku: Sloučení vybraných termínů z aplikačního a doménového slovníku.
- Modelování aplikačních scénářů: Doplnění zpracovaných diagramů případů užití o UML diagramy aktivit a diagramy tříd.
- Vytvoření referenčního výkladového slovníku: Zpracování neformálních definic v přirozeném jazyce pro termíny zahrnuté v referenčním slovníku.

Fáze Návrh – cílem je vypracování ontologické struktury pro položky referenčního výkladového slovníku. Vstupem je referenční výkladový slovník a diagramy tříd a aktivit. Výstupem je sémantická síť a její převod do formátu ontologie.

- Modelování pojmů: Pro kategorizaci pojmů autoři nabízejí vlastní metodiku OPAL²⁵. V průběhu konceptualizace se pojmy rozčlení do kategorií (aktéři, objekty, procesy, zprávy a atributy).
- Modelování hierarchických a doménově specifických vztahů: Při definování hierarchických vztahů je doporučen postup *middle-out* od nejméně frekventovaných pojmů k obecným a k specifickým. Základní generická taxonomie se doplní o další typy vztahů – agregace a asociace. Výsledná sémantická síť je graficky znázorněna UML diagramem tříd.

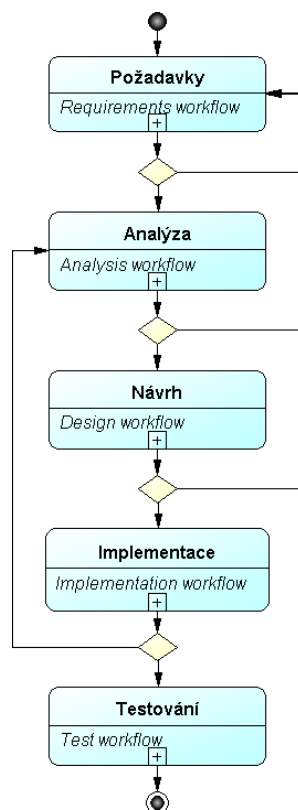
Fáze Implementace – cílem je zakódování ontologie do rigorózního formálního jazyka. Výstupem je ontologie v jazyce OWL.

- Výběr formálního jazyka: Zvažuje se vyjadřovací síla, výpočetní složitost usuzovací metody a akceptování jazyka v rámci dané komunity. Autoři doporučují jazyk OWL.
- Formalizace ontologie: Do formálního jazyka se převede množina definovaných pojmů, vztahy mezi pojmy a množina sémantických axiomů.

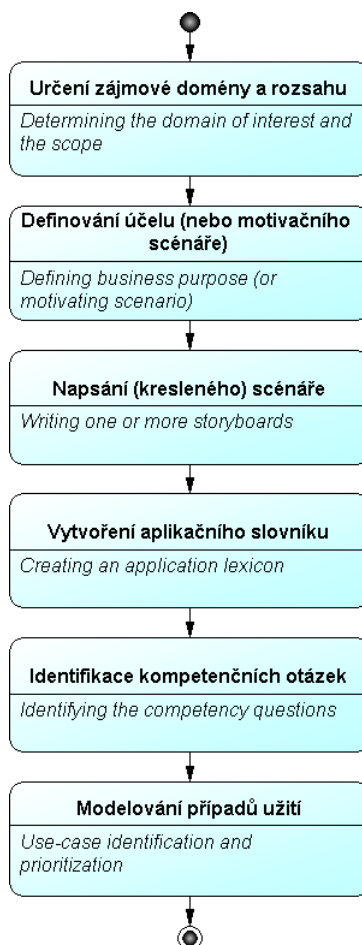
Fáze Testování – cílem je kontrola splnění požadavků na ontologii a posouzení sémantické a pragmatické kvality (užitečnosti pro uživatele) ontologie.

- Kontrola konzistence: Využití automatických kontrol provedených specializovaným softwarem (tzv. *reasoner*, např. [Racer](#), [Pellet](#)), kontrolujícím správnost definovaných logických axiomů a pravidel (např. disjunktnost hierarchie).
- Verifikace pokrytí: Testuje se, zda je možné sémanticky anotovat zdroj spadající do aplikační domény s použitím pojmů obsažených v ontologii.
- Odpovědi na kompetenční otázky: Testuje se, zda je možné kompetenční otázky, zformulované ve fázi Požadavky, zodpovědět prostřednictvím obsahu ontologie.

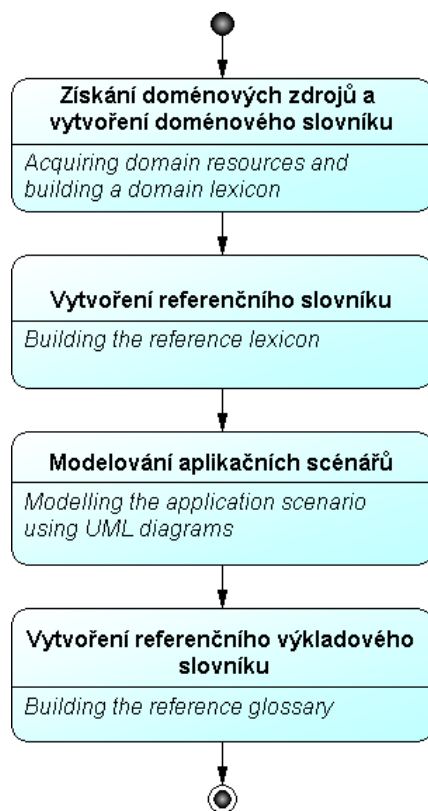
Procesní grafy (obrázky 13 – 18):



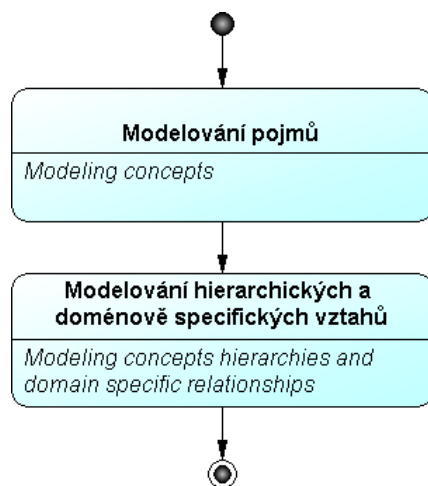
Obr. 13 UPON – členění na základní fáze



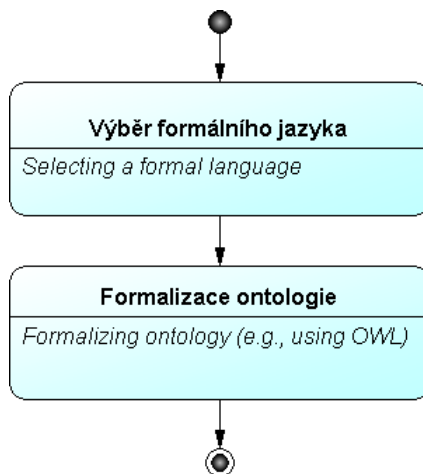
Obr. 14 UPON – detailní členění fáze Požadavky



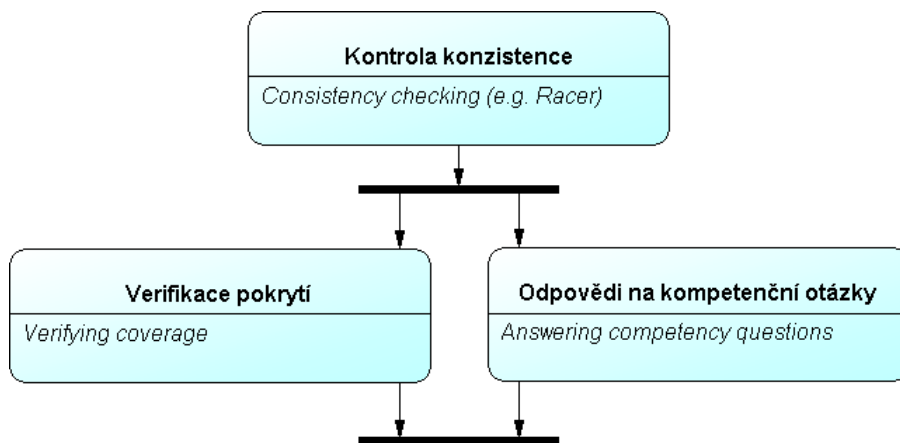
Obr. 15 UPON – detailní členění fáze Analýza



Obr. 16 UPON – detailní členění fáze Návrh



Obr. 17 UPON – detailní členění fáze Implementace



Obr. 18 UPON – detailní členění fáze Testování

Zdroje:

DE NICOLA, Antonio, MISSIKOFF, Michele a NAVIGLI, Roberto. A proposal for a Unified Process for Ontology Building: UPON. In *DEXA 2005*. Berlin: Springer, 2005, s. 655-664. Lecture Notes in Computer Science, no. 3588. ISSN 0302-9743

DE NICOLA, Antonio, MISSIKOFF, Michele a NAVIGLI, Roberto. A software engineering approach to ontology building. *Information systems*. 2009, vol. 34, no. 2 (April), s. 258-275. ISSN 0306-4379. DOI 10.1016/j.is.2008.07.002

3.6 Metodika tvorby tezauru

Určení, specifika: Na tvorbě metodik pro vytvoření tezaurů se od 70. let 20. století podílelo mnoho zástupců odborné veřejnosti a četné významné instituce: [UNESCO](#), [IFLA](#), [ISO](#) (konkrétně technická komise [ISO/TC 46](#) Information and documentation), národní standardizační instituce [BSI](#) a [ANSI / NISO](#), na národní úrovni se věnovali už v rané fázi vývoje problematice tvorby tezaurů i odborníci z tehdejšího Československa^{26, 27, 28, 29}. Autoritativním manuálem se stala příručka *Thesaurus construction and use*, kterou autoři Jean Aitchison, Alan Gilchrist a David Bawden v roce 2000 vydali již ve 4. vydání. Výstupy těchto aktivit v podobě publikovaných dokumentů mapuje v historickém přehledu tabulka 2. Jak naznačují postupné změny názvů, od původního striktního zaměření na návrh tezaurů se záběr rozšiřuje směrem k obecnějšímu pojetí, použitelnému pro širší spektrum typů řízených slovníků a jiných selekčních jazyků. Od roku 2008 připravuje subkomise [SC 9 ISO/TC 46](#) Identification and description na základě britské normy BS 8723 novou verzi standardů ISO 2788:1986 a ISO 5964:1985 jako dvoudílnou normu číslo 25964^{30, 31}. Očekává se, že tato norma přinese sjednocující pohled na pravidla a metody návrhu tezaurů a bude reflektovat změny, k nimž došlo zavedením informačních technologií do této oblasti (softwarové aplikace pro tvorbu a využívání tezaurů, technologie fulltextového vyhledávání ad.)

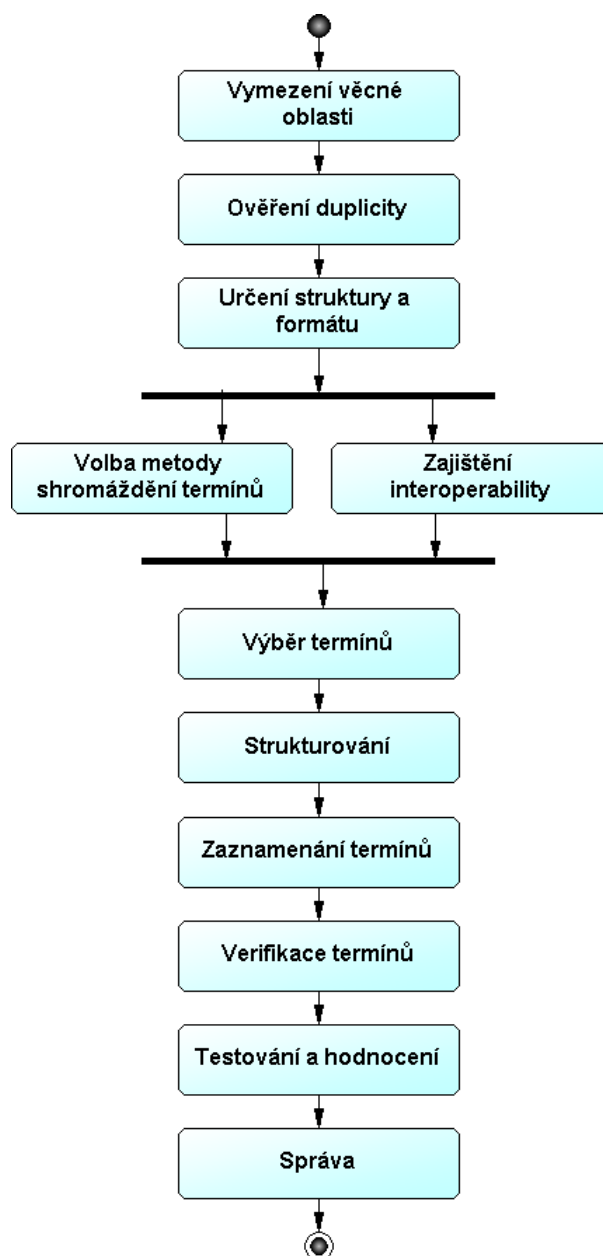
Rok	Producent / autor	Typ a předmět normy	Vydání / verze
1971	UNESCO	Jednojazyčné tezaury ³²	1
1972	Jean AITCHISON, Alan GILCHRIST	Tezaury – příručka ³³	1
1973	UNESCO	Jednojazyčné tezaury ³⁴	2
1974	ISO/TC 46	ISO 2788 – jednojazyčné tezaury ³⁵	1
	ANSI	Z39.19 – tezaury ³⁶	1
1976	UNESCO	Vícejazyčné tezaury ³⁷	1
1980	ANSI/NISO	Z39.19 – tezaury	2
1985	ISO/TC 46	ISO 5964 – vícejazyčné tezaury ³⁸	1
1986	ISO/TC 46	ISO 2788 – jednojazyčné tezaury ³⁹	2
1987	Jean AITCHISON, Alan GILCHRIST	Tezaury – příručka ⁴⁰	2
1992	ISO/TC 46	ČSN ISO 5964 – vícejazyčné tezaury ⁴¹ (vydání v češtině)	1
1993	ANSI/NISO	Z39.19 – jednojazyčné tezaury ⁴²	3
1995	ISO/TC 46	ISO 2788 / ČSN 01 0193 – jednojazyčné tezaury ⁴³ (vydání v češtině, obsahově kompatibilní s ISO 2788, s příklady přizpůsobenými specifikům národního jazyka)	2
1997	Jean AITCHISON, Alan GILCHRIST, David BAWDEN	Tezaury – příručka ⁴⁴	3
1999	IFLA	Principy předmětových selekčních jazyků ⁴⁵	1
2000	Jean AITCHISON, Alan GILCHRIST, David BAWDEN	Tezaury – příručka ⁴⁶	4
2005	BSI	BS 8723 – strukturované slovníky, části 1 ⁴⁷ a 2 ⁴⁸	1
	ANSI/NISO	Z39.19 – jednojazyčné řízené slovníky ⁴⁹	4
2007	BSI	BS 8723 – strukturované slovníky, části 3 ⁵⁰ a 4 ⁵¹	1
2008	ISO/TC 46	ISO 25964 – tezaury	v přípravě
2009	IFLA	Vícejazyčné tezaury ⁵²	1

Tab. 2 Historický přehled vývoje metodik pro tvorbu tezaurů

Klíčové fáze:

- Vymezení věcné oblasti: Vymezení obsahového zaměření tezauru. Určení jádra, které bude zpracováno nově a do hloubky, a okrajových oblastí, pro jejichž pokrytí mohou být využity existující tezaury nebo jejich části.
- Ověření duplicity: Průzkum existujících tezurů zpravidla potvrdí, že pro zvolenou specifickou oblast zatím samostatný tezaurus neexistuje, ale v příbuzných oborech se najdou obsahové části, využitelné při tvorbě nového tezauru. Po přijetí rozhodnutí o vytvoření nového tezauru se doporučuje tento záměr zveřejnit.
- Určení struktury a formátu: Definování konstrukční architektury tezauru, založené na stanoveném způsobu jeho používání. Určení úrovně specifičnosti a lexikálního formátu termínů.
- Volba metody shromáždění termínů: Volí se mezi těmito přístupy: *top-down* / *bottom-up*, dedukce / indukce a jejich případná kombinace, a automatizovaná metoda (tj. identifikace kandidátních termínů v textech, sledování frekvence používání termínů při indexování dokumentů nebo v dotazech uživatelů).
- Zajištění interoperability: Výběr metody pro zajištění interoperability navrhovaného tezauru s již existujícími slovníky. Na výběr jsou metody derivace (odvozování z modelového zdroje), překlad / adaptace, satelitní slovníky (obdobné metodě derivace), přímé propojování prostřednictvím jednotlivých termínů, mapování, spojení prostřednictvím propojovacího slovníku.
- Výběr termínů: Zdrojem pro sběr termínů se mohou stát standardizované terminologické zdroje, analýza dotazů, excerpty textů, znalosti uživatelů a expertů a znalosti sestavovatelů tezauru.
- Strukturování: Návrh základních kategorií pro uspořádání vybraných termínů a jejich vzájemných vztahů. Volba pořadacího principu: a) podle oborů nebo tematických oblastí, b) podle faset, c) kombinovaný přístup – např. na nejvyšší úrovni podle tematických oblastí a v jejich rámci podle faset.
- Zaznamenání termínů: Pro každý termín se pořídí samostatný záznam, v němž se popíší jeho vlastnosti a vztahy k ostatním termínům (ekvivalence, hierarchie, asociace). Tato fáze tedy zahrnuje i strukturování tezauru – analýzu a seskupování termínů v rámci stanovených kategorií.
- Verifikace termínů: Ověření správnosti termínů vzhledem k stanoveným pravidlům výběru. Porovnání věcného rozsahu a formy termínů s autoritativními zdroji (odborné slovníky, encyklopedie, existující tezaury a klasifikační schémata).
- Testování a hodnocení: Doporučuje se realizovat tuto fázi před zveřejněním tezauru a v pravidelných intervalech v průběhu jeho používání.
- Správa: Aktualizace a doplňování slovníku – přidávání nových termínů, modifikace existujících termínů, odstraňování nadužívaných nebo nepoužívaných termínů. Doporučuje se formalizovat návrhy na aktualizaci pomocí formulářů a veškeré provedené změny zaznamenávat v historických poznámkách, případně v odkazech.

Procesní graf (obrázek 19):



Obr. 19 Metodika tvorby tezauru

Zdroje:

Přes množství publikovaných dokumentů (viz tab. 2) nebylo pro naše účely možné žádný z nich vybrat jako referenční zdroj – v podstatě všechny texty jsou primárně orientovány na stanovení pravidel a metodiky tvorby tezauru jsou v nich obsaženy implicitně, bez podrobného definování jednotlivých procesů a jejich návazností. Popis této metodiky v našem přehledu je tedy zpracován odlišným způsobem oproti ostatním metodikám, je výsledkem kompilace s využitím více zdrojů⁵³.

3.7 Denton – Spiteri

Autoři: William Denton (York University of Toronto), Louise F. Spiteri (Dalhousie University, Halifax)

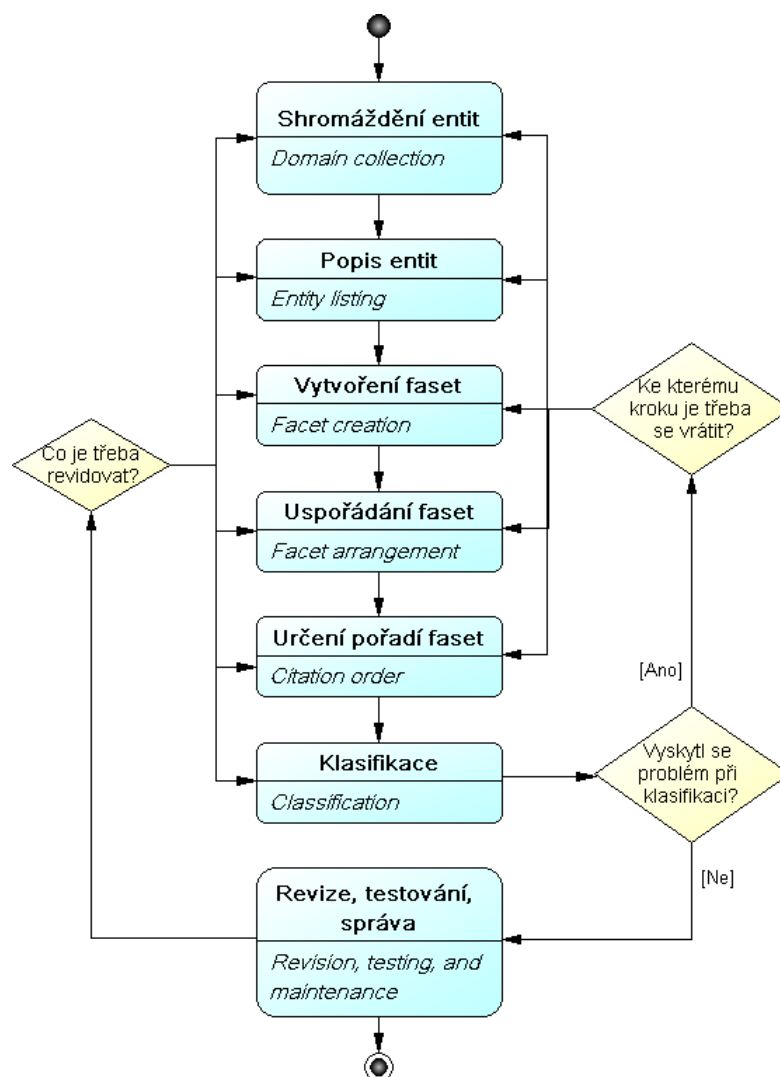
Rok vzniku / publikování: 1998 (Spiteri) / 2003 (Denton)

Určení, specifika: Metodika tvorby fasetové klasifikace pro publikování webového obsahu, určená pro malé a středně velké kolekce dokumentů. Východiskem jsou čtyři fáze tvorby fasetové klasifikace, jež v roce 1960 definoval Brian Campbell Vickery⁵⁴. V metodice navržené Williamem Dentonem jsou uplatněny v pozmeněném pořadí a doplněny o další tři fáze. Ústřední fáze – vytvoření a uspořádání faset, jsou založeny na zjednodušeném modelu fasetové analýzy, který vypracovala Louise Spiteri kompilací vybraných kánonů, postulátů a principů S. R. Ranganathana a principů fasetové analýzy zpracovaných Skupinou pro výzkum klasifikace (Classification Research Group – CRG). Ve vodopádovém životním cyklu se uplatňují četné prvky iterace, v jednotlivých fázích se střídavě užívají postupy *top-down* a *bottom-up*.

Klíčové fáze:

- Shromáždění entit: Shromáždění reprezentativního vzorku entit pro danou doménu; u malých kolekcí je doporučeno zahrnout entity z celé domény.
- Popis entit: Vytvoření popisu jednotlivých entit. Sémantický rozklad popisů – rozdělení vět na části, přeuspořádání slov, rozpoznání základních pojmů v jednotlivých částech vět a jejich izolování.
- Vytvoření faset: Prozkoumání výsledných termínů a nalezení obecných kategorií na nejvyšší úrovni, vyskytujících se ve všech entitách. Omezení nalezených kategorií na množinu vzájemně se vylučujících a ve svém celku vyčerpávajících faset, použitelných pro všechny termíny vytvořené v předchozí fázi. Principy pro výběr faset:
 - diferenciace – použité kritérium členění má vést k jasnému rozlišení jednotlivých částí
 - relevance – zvolené fasety mají odpovídat účelu, předmětu a rozsahu klasifikace
 - průkaznost (*ascertainability*) a permanence – fasety mají reprezentovat trvalé a měřitelné charakteristiky klasifikovaných objektů
 - homogenita a disjunktnost – každá faseta smí reprezentovat pouze jednu charakteristiku členění
 - základní kategorie mají být přizpůsobeny specifickým vlastnostem a kontextu klasifikovaných objektů
- Uspořádání faset: Naplnění faset termíny vytvořenými ve fázi Popis entit, tak aby bylo možné zatřídit všechny shromážděné entity. Uspořádání termínů uvnitř fasety podle zvoleného principu (abecední, hierarchický, chronologický ad.). Volba preferovaných termínů, tvorba řízeného slovníku.
- Určení pořadí faset: Rozhodnutí o standardním (defaultním) uvádění pořadí faset (např. na domovské stránce).
- Klasifikace: Zatřídění celého obsahu domény. Analýza všech entit za použití faset a použití správných termínů pro popis každé entity.
- Revize, testování, správa: Při jakýchkoli problémech zjištěných v etapě Klasifikace následuje iterace – návrat k předchozí fázi, jež je relevantní k vyřešení problému. Následuje uživatelské testování a pravidelná správa klasifikace (změna terminologie, nové entity, změny v doméně).

Procesní graf (obrázek 20):



Obr. 20 Denton – Spiteri

Zdroje:

DENTON, William. *How to make faceted classification and put it on the web* [online]. Miskatonic University Press, November 2003 [cit. 19.4.2011]. Dostupné z: <http://www.miskatonic.org/library/facet-web-howto.html>

SPITERI, Louise F. A simplified model for facet analysis: Ranganathan 101. *Canadian journal of information and library science*. 1998, vol. 23, no. 1-2 (April-July), s. 1-30. ISSN 1195-096X. Dostupné též z: http://iainstitute.org/en/learn/research/a_simplified_model_for_facet_analysis.php [cit. 19.4.2011]

		Grűninger – Fox	Ushold – King	Noy – McGuinness	Methontology	UPON	Metodika tezauru	Denton – Spiteri
		1995	1995	2001	1997	2005	1971–2009	1998/2003
Zdroj	umělá inteligence	●	●	●				
	softwarové inženýrství				●	●		
	knihovní a informační věda						●	●
Využití softwaru	intelektuální	●	●	●			●	●
	poloautomatizovaná				●	●		
	automatizovaná							
Postup řešení	top-down			●			●	●
	middle-out	●	●	●	●	●	●	
	bottom-up			●			●	●
Životní cyklus	vodopádový	●	●				●	●
	iterativní	●		●	●	●		●
Účel metodiky	plánovací	●	●		●	●		●
	konstrukční			●	●	●	●	
Formát	technický standard				●	●	●	
	odborný text	●	●	●				●

Tab. 3 Typologie metodik

Literatura – přehledy metodik ontologického inženýrství

BREITMAN, Karin Koogan, CASANOVA, Marco Antonio a TRUSZKOWSKI, Walter. *Semantic web: concepts, technologies and applications*. 1. ed. London: Springer, 2007. 327 s. NASA Monographs in Systems and Software Engineering. ISSN 1860-0131. ISBN-10 1-84628-581-X, ISBN-13 978-1-84628-581-3, e-ISBN 978-1-84628-710-7. Kapitola 8, Methods for ontology development, s. 155-173.

Popisované metodiky: Grüninger – Fox / TOVE, Horrocks ontology development method, KACTUS, LEL, Methontology, Noy – McGuinness, Uschold – King / Enterprise.

FERNÁNDEZ-LÓPEZ, Mariano. Overview of methodologies for building ontologies. 13 s. In *Proceedings of IJCAI-99 Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends*. Held on August 2, 1999 in conjunction with the Sixteenth International Joint Conference on Artificial Intelligence. City Conference Center, Stockholm, Sweden. ISBN 90-5470-085-8. Dostupné z: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/> [cit. 24.4.2011]

Popisované metodiky: Grüninger – Fox / TOVE, KACTUS, Methontology, SENSUS, Uschold – King / Enterprise.

GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano a CORCHO, Oscar. *Ontological Engineering: With examples from the areas of knowledge management, e-commerce and the Semantic Web*. 1. ed. London: Springer, 2004. Kapitola 3, Methodologies and Methods for Building Ontologies, s. 107–197.

Popisované metodiky jsou rozděleny na 2 skupiny: 1) “klasické” metodiky tvorby ontologie “na zelené louce” nebo znovupoužitím existujících ontologií: Cyc method, Grüninger – Fox / TOVE, KACTUS, Methontology, On-to-knowledge, SENSUS, Uschold – King / Enterprise. Tyto metodiky jsou porovnány podle 13 jednotně stanovených kritérií. 2) speciální metodiky věnované určitému aspektu vývoje ontologií: metodika reengineeringu ontologií (adaptovaná metodika Methontology), metodiky učení ontologií (Maedche et al., Aussenac-Gilles et al.), metodiky slučování ontologií (ONIONS, FCA-Merge, PROMPT), metodika kooperativní tvorby ontologií (Co4), metodika hodnocení ontologií (OntoClean).

JONES, Dean, BENCH-CAPON, Trevor a VISSER, Pepijn. Methodologies for ontology development. In *IT & KNOWS – Informations Technology And Knowledge Systems. Proceedings of the 15th IFIP World computation congress*, Budapest, 1998. 1998, s. 62-75. ISBN 3-85403-122-X. DOI 10.1.1.52.2437

Popisované metodiky jsou rozděleny na 2 skupiny: 1) metodiky pokrývající celý vývojový cyklus – Grüninger – Fox / TOVE, KBSI IDEF5, Methontology, Uschold – King / Enterprise, 2) metodiky věnované určitému dílčímu aspektu vývoje ontologií – CommonKADS / KACTUS, Guarino et al., MENELAS, Mikrokosmos, ONIONS, Ontolingua, PHYSSYS, PLINIUS, SENSUS. V závěru autoři formulují 5 obecných problémů, které jsou řešeny všemi metodikami.

MIZOGUCHI, Riichiro. Tutorial on ontological engineering. Part 2: Ontology Development, Tools and Languages. *New Generation Computing*. 2004, vol. 22, no. 1 (January), s. 61-96. ISSN 0288-3635 (Print), 1882-7055 (Online)

Popisované metodiky: AFM: Activity-first method, Grüninger – Fox / TOVE, Methontology, On-to-knowledge, Uschold – King / Enterprise.

OntoWeb. Deliverable 1.4: A survey on methodologies for developing, maintaining, evaluating and reengineering ontologies [online]. Version 1.0. Ed. Mariano Fernández-López. 12. 12. 2002 [cit. 19.4.2011]. 56 s. Dostupné z: http://www.kde.cs.uni-kassel.de/stumme/papers/2002/OntoWeb_Del_1-4.pdf

Popisované metodiky: Co4, Cyc method, FCA-Merge, Grüninger – Fox / TOVE, IF-Map, KACTUS, (KA)2, Methontology, metoda reengineeringu ontologií skupiny Ontology Group UPM, MOMIS, ONIONS, OntoClean, On-to-knowledge, PROMPT, SENSUS, Uschold – King / Enterprise.

PINTO, Helena Sofia a MARTINS, João P. Ontologies: how can they be built? *Knowledge and Information Systems*. 2004, vol. 6, no. 4 (July), s. 441–464. ISSN 0219-1377 (printed version), ISSN 0219-3116 (electronic version). DOI 10.1007/s10115-003-0138-1

Popisované metodiky: Grüninger – Fox / TOVE, Methontology, OntoClean, Uschold – King / Enterprise.

Poznámky a použité zdroje

¹ *Ontology definition metamodel (ODM)* [online]. Version 1.0. OMG Document Number: formal/2009-05-01. Object Management Group, May 2009 [cit. 19.4.2011], s. 31. Dostupné z: <http://www.omg.org/spec/ODM/1.0>. Zvýraznění části textu H.K.

² Je vhodné upozornit, že „realita“ v tomto pojetí zahrnuje nejen to, co aktuálně a prokazatelně existuje, ale i představy a myšlenkové konstrukce.

³ „Jsou to všechny entity, o které se zajímáme, které byly, jsou nebo mohou být.“ ČSN ISO/TR 9007 (97 9702). *Systémy zpracování informací. Pojmy a terminologie pro pojmové schéma a informační základnu*. Praha: Český normalizační institut, 1995, s. 16. Aplikací tohoto obecného pojmu jsou potom nejrůznější doménová univerza, např. bibliografické univerzum, informační univerzum apod.

⁴ I když i takové snahy v podobě „základních“, „generických“, „upper“, „top level“ ontologií je možné zaznamenat – příkladem jsou např. BFO, Cyc, KR ontologie, ale i lexikální ontologie Wordnet. Tyto základní ontologie jsou však spíše zamýšleny jako zdroje pro tvorbu specializovaných doménových ontologií, než pro přímé využití při organizaci a komunikaci znalostí.

⁵ BUCHALCEVOVÁ, Alena. *Metodiky budování informačních systémů*. Praha: Oeconomica, 2009, s. 95-119.

⁶ GÓMEZ-PÉREZ, Asunción, FERNÁNDEZ-LÓPEZ, Mariano a CORCHO, Oscar. *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic web*. 1. ed. London: Springer, 2004. 403 s. ISBN 978-1-85233-551-9

⁷ ALLEMANG, Dean a HENDLER, James. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. Morgan Kaufmann, 2008. 352 s. ISBN 978-0-12-373556-0

⁸ STAAB, Steffen a STUDER, Rudi (eds.) *Handbook on ontologies*. 2. ed. Dordrecht: Springer, 2009. 811 s. ISBN 978-3-540-70999-1, e-ISBN 978-3-540-92673-3. DOI 10.1007/978-3-540-92673-3

⁹ HEPP, Martin, LEENHEER, Pieter de, MOOR, Aldo de, SURE, York (eds.) *Ontology Management: Semantic Web, Semantic Web Services, and Business Applications*. Springer, 2008. 295 s. Semantic Web and Beyond, vol. 7. ISBN 978-0-387-69899-1

-
- ¹⁰ SVÁTEK, Vojtěch. Ontologie a WWW. In *DATAKON 2002: sborník databázové konference*. Brno, Česká republika 19.-22. října 2002. Brno: Masarykova univerzita, 2002, s. 27-55.
- ¹¹ SVÁTEK, Vojtěch a VACURA, Miroslav. Ontologické inženýrství. In *DATAKON 2007: sborník databázové konference*. Brno, Česká republika 20.-23. října 2007. Brno: Masarykova univerzita, 2007, s. 60-91. Dostupné též z: <http://nb.vse.cz/~svatek/dkon07final.pdf> [cit. 19.4.2011]
- ¹² Terminologická poznámka: Dalším používaným termínem je metodologie – ten v češtině označuje vědní obor zkoumající metody. Anglický ekvivalent „*methodology*“ však má dva významy: 1) vědní obor zkoumající metody, 2) metodika. Nesprávným překladem z angličtiny se tak v českých textech lze setkat s termínem metodologie i tam, kde jde o metodiku.
- ¹³ FOWLER, Martin. *Destilované UML*. 1. vyd. Praha: Grada, 2009. Kapitola 2, Proces vývoje, s. 37-49. ISBN 978-80-247-2062-3
- ¹⁴ ROYCE, Winston W. Managing the Development of Large Software Systems. *Proceedings of IEEE WESCON*, 1970, no. 26 (August), s. 1–9. Dostupné z: <http://www.cs.umd.edu/class/spring2003/cmsc838p/Process/waterfall.pdf> [cit. 24.4.2011]
- ¹⁵ BOEHM, Barry. A Spiral Model of Software Development and Enhancement. *IEEE Computer*. 1988, vol. 21, no. 5 (May), s. 61-72. ISSN 0018-9162. Dostupné z: <http://csse.usc.edu/csse/TECHRPTS/1988/usccse88-500/usccse88-500.pdf> [cit. 24.4.2011]
- ¹⁶ MINSKY, Marvin. *A framework for representing knowledge* [online]. Cambridge: MIT Media Laboratory, 1974 [cit. 19.4.2011]. MIT-AI Laboratory Memo 306, June, 1974. Dostupné z: <http://web.media.mit.edu/~minsky/papers/Frames/frames.html> [cit. 24.4.2011]
- ¹⁷ PRISS, Uta. Formal concept analysis in information science. In *Annual review of information science and technology*. Vol. 40, 2006. Ed. Blaise Cronin. Medford: Information Today for American Society for Information Science & Technology (ASIST), 2006, s. 521-543. ISBN 1-57387-242-3; ISSN 0066-4200
- ¹⁸ BERKA, Petr, MAŘÍK, Vladimír a SVÁTEK, Vojtěch. Proces získávání znalostí. In MAŘÍK, Vladimír, ŠTĚPÁNKOVÁ, Olga a LAŽANSKÝ, Jiří. *Umělá inteligence 2*. 1. vyd. Praha: Academia, 1997, s. 107-124. ISBN 80-200-0504-8
- ¹⁹ „...[O]ntologie má konceptuálně velmi blízko k fasetové struktuře a ... fasetová analýza má hodně co nabídnout jako metoda pro tvorbu ontologií.“ BROUGHTON, Vanda. The need for faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings: New Information Perspectives*. 2006, vol. 58, no. 1/2, s. 66. ISSN 0001-253X. DOI 10.1108/00012530610648671
- ²⁰ SCHWARZ, Josef. Současný stav a trendy automatické indexace dokumentů: přehledová studie. *Ikaros* [online]. 2003, roč. 7, č. 3 [cit. 19.4.2011]. ISSN 1212-5075. Dostupný z: <http://www.ikaros.cz/node/1300>
- ²¹ IEEE 1074-1995. *IEEE Standard for developing software life cycle processes*. Institute of Electrical and Electronics Engineers. 1996. 96 s. ISBN 1-55937-588-4. DOI 10.1109/IEEESTD.1996.79663
- ²² JACOBSON, Ivar, BOOCH, Grady a RUMBAUGH, James. *The Unified Software Development Process*. Reading: Addison-Wesley Professional, 1999. 512 s. ISBN 978-0201571691
- ²³ Zdroj: DE NICOLA, Antonio, MISSIKOFF, Michele a NAVIGLI, Roberto. A software engineering approach to ontology building. *Information systems*. 2009, vol. 34, no. 2 (April), s. 260. ISSN 0306-4379. DOI 10.1016/j.is.2008.07.002
- ²⁴ Dohoda uživatelů ontologie o významu používaných znaků (slov, termínů) pro vyjádření pojmů a jejich vztahů.

-
- ²⁵ D'ANTONIO, Fulvio, MISSIKOFF, Michele a TAGLINO, Francesco. Formalizing the OPAL eBusiness ontology design patterns with OWL. In *Enterprise interoperability II: new challenges and approaches*. Proceedings of the Third International conference on interoperability for enterprise applications and software, I-ESA 2007. London: Springer, 2007, s. 345-356. ISBN 978-1-84628-858-6. DOI 10.1007/978-1-84628-858-6_38
- ²⁶ SECHSER, Otto, et al. *Studie metod stavby tezurů*. 1. vyd. Praha: UVTEI, 1967. 119 s., 33 s. obr. příl. Materiály k metodice a technice informací, sv. 1.
- ²⁷ TOMAN, Jiří. *Postup při sestavování tezauru*. Praha: Státní technická knihovna, 1968. 31 s. Metodický leták č. 68.
- ²⁸ MOJŽÍŠEK, Josef. Metodika tvorby tezauru: Závěrečná zpráva výzkumného úkolu P-18-121-001-04. 1. vyd. Praha: ÚVTEI, 1974. 142, [1] s.
- ²⁹ VORÁČEK, Josef. *Tvorba tezauru v českém jazyce*. Praha: UVTEI, 1974. 51 s. Metodické letáky, sv. 101.
- ³⁰ ISO/FDIS 25964-1. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval*. 144 s. Status: Under development. Stage: 50.20 (2011-04-06)
- ³¹ ISO/CD 25964-2. *Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies*. Status: Under development. Stage: 30.20 (2011-03-01)
- ³² UNESCO. *Guidelines for the establishment and development of monolingual thesauri for information retrieval*. 1. revised ed. Paris: UNESCO, 1971 [cit. 24.4.2011]. 30 s. SC/WS/500. Revised version of document UNESCO/SC/MD 20 Guidelines for the establishment and development of monolingual scientific and technical thesauri for information retrieval. Dostupné z: <http://unesdoc.unesco.org/images/0000/000059/005951EB.pdf>
- ³³ AITCHISON, Jean a GILCHRIST, Alan. *Thesaurus construction: a practical manual*. 1. ed. London: Aslib, 1972. 95 s. ISBN 0851420427
- ³⁴ UNESCO. *UNISIST Guidelines for the establishment and development of monolingual thesauri*. 2. ed. Paris: UNESCO, 1973 [cit. 24.4.2011]. 37 s. SC/WS/555. Dostupné z: <http://unesdoc.unesco.org/images/0000/000059/005951EB.pdf>
- ³⁵ ISO 2788:1974. *Documentation – Guidelines for the establishment and development of monolingual thesauri*. 1. ed. Geneva: International Organization for Standardization, 1974.
- ³⁶ Z39.19-1974. *American national standard guidelines for thesaurus structure, construction and use*. 1. ed. New York: American National Standards Institute, 1974.
- ³⁷ UNESCO. *UNISIST Guidelines for the establishment and development of multilingual thesauri*. 1. ed. Paris: UNESCO, 1976 [cit. 24.4.2011]. 49 s. SC.76/WS/102. Dostupné z: <http://unesdoc.unesco.org/images/0002/000205/020561EB.pdf>
- ³⁸ ISO 5964:1985. *Documentation – Guidelines for the establishment and development of multilingual thesauri*. 1. ed. Geneva: International Organization for Standardization, 1985. 61 s.
- ³⁹ ISO 2788:1986. *Documentation – Guidelines for the establishment and development of monolingual thesauri*. 2. ed. Geneva: International Organization for Standardization, 1986. 32 s.
- ⁴⁰ AITCHISON, Jean a GILCHRIST, Alan. *Thesaurus construction: a practical manual*. 2. ed. London: Aslib, 1987 (1990 printing). 173 s. ISBN 0851421970
- ⁴¹ ČSN ISO 5964 (01 0172). *Pokyny pro vypracování a rozvíjení vícejazyčných tezurů*. Praha: Český normalizační institut, 1992. 60 s.

-
- ⁴² Z39.19-1993. *Guidelines for the construction, format and management of monolingual thesauri*. 3. ed. Bethesda (MD): American National Standards Institute, 1993.
- ⁴³ ČSN 01 0193. Dokumentace. Pokyny pro vypracování a rozvíjení jednojazyčných tezaurů. Praha: Český normalizační institut, 1995. 52 s.
- ⁴⁴ AITCHISON, Jean, GILCHRIST, Alan a BAWDEN, David. *Thesaurus construction and use: a practical manual*. 3. ed. London: Aslib, 1997. 212 s. ISBN 0-85142-390-6
- ⁴⁵ *Principles Underlying Subject Heading Languages (SHLs)*. Ed. by Lopes, Maria Ines, Beall, Julianne, Working Group on Principles Underlying Subject Heading Languages. In collab. with Standing Committee of the IFLA Section on Classification and Indexing. De Gruyter Saur, 1999. 183 s. Series: UBCIM Publications. New Series 21. ISBN 978-3-598-11397-0.
- ⁴⁶ AITCHISON, Jean, GILCHRIST, Alan a BAWDEN, David. *Thesaurus construction and use: a practical manual*. 4. ed. Chicago: Fitzroy Dearborn Publishers, 2000. 218 s. ISBN 1579582737
- ⁴⁷ BS 8723-1:2005. *Structured vocabularies for information retrieval. Guide. Definitions, symbols and abbreviations*. London: British Standards Institution, 2005. 14 s. ISBN 0-580-46798-8
- ⁴⁸ BS 8723-2:2005. *Structured vocabularies for information retrieval. Guide. Thesauri*. London: British Standards Institution, 2005. 64 s. ISBN 0-580-46799-6
- ⁴⁹ ANSI/NISO Z39.19-2005. *Guidelines for the construction, format, and management of monolingual controlled vocabularies* [online]. 4. ed. Bethesda (MD): American National Standards Institute, 25. 7. 2005 [cit. 19.4.2011]. 184 s. ISSN 1041-5653. ISBN 1-880124-65-3. Dostupné z: <http://www.niso.org/kst/reports/standards/>
- ⁵⁰ BS 8723-3:2007. *Structured vocabularies for information retrieval. Guide. Vocabularies other than thesauri*. London: British Standards Institution, 2007. 52 s. ISBN 978-0-580-63072-9
- ⁵¹ BS 8723-4:2007. *Structured vocabularies for information retrieval. Guide. Interoperability between vocabularies*. London: British Standards Institution, 2007. 62 s. ISBN 978-0-580-63073-6
- ⁵² IFLA. *Guidelines for Multilingual Thesauri* [online]. The Hague: IFLA Working Group on Guidelines for Multilingual Thesauri Classification and Indexing Section, 2009 [cit. 19.4.2011]. 26 s. IFLA Professional Reports, No. 115. ISBN 978-90-77897-35-5. ISSN 0168-1931. Dostupné z: <http://archive.ifla.org/VII/s29/pubs/Profrep115.pdf>
- ⁵³ Vycházíme zejména z těchto zdrojů: AITCHISON, Jean, GILCHRIST, Alan a BAWDEN, David. *Thesaurus construction and use: a practical manual*. 3. ed. London: Aslib, 1997. 212 s. ISBN 0-85142-390-6; ČSN ISO 5964 (01 0172). *Pokyny pro vypracování a rozvíjení vícejazyčných tezaurů*. Praha: Český normalizační institut, 1992. 60 s.; ČSN 01 0193. *Dokumentace. Pokyny pro vypracování a rozvíjení jednojazyčných tezaurů*. Praha: Český normalizační institut, 1995. 52 s.
- ⁵⁴ Jedná se o následující fáze: 1. Roztřídění termínů v dané oblasti do homogenních, navzájem se vylučujících faset, z nichž každá je odvozena z nadřazeného (*parent*) univerza podle jednoho kritéria členění (s. 12); 2. Určení pořadí, v němž budou fasety použity ke konstrukci složených hesel; 3. Vytvoření takové notace pro dané schéma, jež dovolí plně flexibilní kombinaci termínů a jež rozmístí předměty do preferovaného pořadí (*filling order*); 4. Použití fasetového schématu takovým způsobem, že umožní jak specifické odkazy, tak požadovanou úroveň obecného průzkumu (s. 13). VICKERY, B. C. *Faceted classification: a guide to construction and use of special schemes*. London: Aslib, 1960, s. 12-13. Cit. dle W. Dentona.